

Open Research Online

The Open University's repository of research publications
and other research outputs

Somatic hypermutation of immunoglobulin light chain genes

Thesis

How to cite:

Goyenechea Corzo, Beatriz (1996). Somatic hypermutation of immunoglobulin light chain genes. PhD thesis. The Open University.

For guidance on citations see [FAQs](#).

© 1996 Beatriz Goyenechea Corzo

Version: Version of Record

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

UNRESTRICTED

Somatic hypermutation of immunoglobulin κ light chain genes

Beatriz Goyenechea Corzo

A thesis submitted in partial fulfilment of the requirements of the Open University in
candidature for the degree of Doctor of Philosophy

June 1996

Medical Research Council
Laboratory of Molecular Biology
Cambridge

Date of submission: June 1996
Date of award: 15 August 1996

ProQuest Number: C534604

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest C534604

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Somatic hypermutation of immunoglobulin genes is a tightly controlled, site directed process. The nucleotide changes are targeted in the region surrounding the rearranged V gene segment of the heavy and light chain loci. The targeting of somatic mutation does neither require V region including the V κ promoter specific features nor a specific distance between the C κ and the 3' enhancer. On the other hand, removal of either the κ 3' enhancer or the κ intron enhancer/MAR causes a drastic reduction in hypermutation, which is particularly marked in the latter case. In the work described in this dissertation I have used transgenic mice carrying immunoglobulin κ transgenes with deleted regulatory elements, to make a tighter delineation of which DNA regions within the κ locus are important for hypermutation. In addition I have used transgenic mice to study the role of the primary sequence and the DNA secondary structure in this process.

This thesis describes the generation of artificial Ig genes constructs to make transgenic lines. κ transgenes were characterized by Southern blot and by ELISA in the seven mouse lines obtained. The role of the 3' enhancer core in the transcription of the κ chain was also studied in stable transfected S107 mouse myeloma cells.

The analysis of the transgenes carrying deletions of regulatory elements revealed that hypermutation was impaired (but not severely affected) when the intron enhancer was deleted. This also was the case when the flanking regions of the 3' enhancer core were deleted. Hypermutation was severely affected (but not abolished) when the core of the 3' enhancer was removed. This reflects the relative roles of both components of the 3' enhancer region in gene expression. The matrix attachment region seems to be essential, since its deletion abolished hypermutation, but not expression, since abundant κ mRNA was still produced. Furthermore these studies showed that the 3' enhancer can drive hypermutation from a position upstream of the V κ region.

The major role of the primary sequence in the formation of hotspots and the mutational machinery preferences for AGY encoded Ser and the bias against TCA encoded Ser as target for hypermutation were disclosed with the analysis of the transgenes carrying V κ Ox1 gene with modifications in the CDR1 sequence. However, whilst hairpin loops do not define which base is going to be targeted, they seem to have a role in the accessibility of the 'favoured' nucleotides for hypermutation.

Declaration

The work described in this thesis was carried out under the supervision of Dr. César Milstein at the Medical Research Council, Laboratory of Molecular Biology, Cambridge, between October 1992 and June 1996. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration unless otherwise stated. I declare that this thesis has not been submitted either in whole or in part to any other academic institution.

Acknowledgements

During the course of this work I have received help and advice from many members of the Laboratory of Molecular Biology. In particular I would like to thank my supervisor, Dr. César Milstein, for his supervision, inspirational ideas and encouragement, especially during the more difficult times. I am also grateful to TECHNE Ltd. and the Association for International Cancer Research for their financial support during the course of this work.

I am very grateful to Richard Pannell for the generation of the hybridomas and his advice regarding many experimental aspects of the tissue culture work. I am also very grateful to Africa González Fernández for her enthusiastic support and helpful discussions, Ermanno Gherardi for his support and encouragement, Cristina Rada for sharing all that she knows about Macs, David Gilmore and Andrew Riddell for all the cell sorting, Walter Gilks for his advice with statistical processing. I would also like to thank Michael Neuberger for criticism and stimulating discussions.

This project relied heavily on the usage of transgenic mice and would not have been possible without the contribution of Tammy Larson and John Jarvis who did the microinjections and Theresa Langford, Gareth King and the staff of the animal facilities at the MRC who looked after the mice.

Thank to all the people who worked in César's group during the period from October 1992 to June 1996 for providing a wonderful working atmosphere.

Finally I would like to thank my husband Lutz for his support and patience.

ABTS	2,2-azino-bis(3-ethylbenzthiazolinesulphonic acid)	MAR	matrix attachment region
bp	base pair(s)	m RNA	messenger RNA
BSA	bovine serum albumin	μ	(immunoglobulin) mu heavy chain
CAT	chloramphenicol acetyl transferase	mAb	monoclonal antibody
cDNA	complementary DNA	MOPS	3-[N-morpholino] propanesulphonic acid
CDR	complementarity determining region	MPW	millipore purified water
Cκ	constant region of the immunoglobulin κ light chain	mw	molecular weight
CSA	chicken serum albumin	nt	nucleotide(s)
CIP	calf intestine phosphatase	LCR	locus controlling region
DMEM	Dulbecco's modified Eagle's medium	λ	(immunoglobulin) lambda light chain
dNTP	deoxynucleotide triphosphate	O.D. _n	optical density at wavelength n
ds	double stranded	o.n.	over night
DTT	dithiothreitol (Cleland's reagent)	PCR	polymerase chain reaction
EDTA	ethylenediaminetetracetic acid (disodium salt)	PE	phycoerythrin
ELISA	enzyme-linked immunosorbent assay	PEG	polyethylene glycol
Ei	intron enhancer	phOx	2-phenyl-oxazol-5-one
EtBr	ethidium bromide	PNA	lectin from <i>Arachis hypogaea</i> (peanut)
FACS	fluorescence activated cell sorting	PP	Peyer's patches
FCS	foetal calf serum	rpm	rotations per minute
FITC	fluorescein isothiocyanate	RT	room temperature
HEPES	N-2-hydroxyethyl piperazine-N-2-ethane sulphonic acid	ss	single stranded
HRP	horse radish peroxidase	SDS	sodium dodecyl sulphate
Ig	immunoglobulin	TEMED	N,N,N',N'-tetramethyl-ethylenediamine
IgH	immunoglobulin heavy chain	Tris	N-Tris-(hydroxymethyl)-methyl-2-aminoethane sulphonic acid
IPTG	isopropyl-β-D-thiogalactopyranoside	UV	ultra violet
κ	(immunoglobulin) kappa light chain	v/v	volume per volume
kb	kilobase pair(s)	w/v	weight per volume
		X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Table of Contents

Abstract	i
Declaration	ii
Acknowledgements	iii
Abbreviations	iv

Chapter One

Introduction

1.1 General background	1
1.1.2 The instructive and selective theories of antibody formation	1
1.1.3 The structure of immunoglobulins	2
1.2 Affinity maturation	3
1.3 Origins of immunoglobulin diversity	3
1.3.1 Germline versus somatic theory	3
1.3.2 Recombination of immunoglobulin genes	4
1.3.3 Diversification of immunoglobulin genes during B cell development	5
1.4. Somatic hypermutation of immunoglobulin genes	5
1.4.1 First indications of somatic hypermutation	5
1.4.2 The rate of somatic hypermutation	6
1.4.3 Target area of somatic hypermutation	6
1.4.4 The timing of somatic hypermutation	6
1.4.5 Intrinsic features of somatic hypermutation	7
1.4.6 Sequences needed to recruit hypermutation	9
1.4.7 Mechanism of somatic hypermutation	9
1.5 Cellular dynamics in affinity maturation	10
1.5.1 The primary response and the germinal centre formation	10
1.5.2 Germinal centre - microenvironment for hypermutation, antigen selection and isotype switching	12
1.5.3 B cell memory	13
1.6 The immune response against 2-phenyl-oxazol-5-one	13
1.6.1 A model for somatic hypermutation	13
1.6.2 Affinity maturation by somatic hypermutation	14
1.6.3 Affinity maturation by repertoire shift	14
1.7 Regulation of the immunoglobulin gene transcription	14
1.7.1 Cis-acting sequences: promoter and enhancers	15
The promoter	15
The intron enhancer	15
The 3' enhancer	16
1.7.2 Regulation of chromatin structure	17
Methylation	17
Matrix attachment regions (MAR)	18

1.7.3 Regulation of the immunoglobulin loci during the B cell development	19
1.8 Methodological approaches for analysing somatic hypermutation	20
1.8.1 The hybridoma approach	20
1.8.2 The PCR approach	21
1.8.3 The transgene approach	21
Analysis of somatic hypermutation in Peyer's patch B cells	21
1.9 Aims of this study	22

Chapter Two

Materials and Methods

2.1 Materials	23
2.1.1 Chemicals and enzymes	23
2.1.2 Standard buffers	24
2.1.3 Media	25
2.1.4 Bacterial strains	25
2.1.5 Origin of plasmids	25
2.1.6 Cell lines	25
2.2 Methods	25
2.2.1 Construction and preparation of plasmids	26
Preparation of plasmid DNA	26
Phenol extraction and ethanol precipitation	26
DNA quantitation	26
Restriction enzyme digest and gel electrophoresis	26
Elution of DNA fragments and ligation reaction	27
Preparation and transformation of competent cells	27
Preparation and electroporation of 'electro-competent cells'	28
Preparation of M13 templates for sequencing	28
2.2.2 Preparation of genomic DNA	29
Isolation of DNA from small numbers of cells	29
2.2.3 Production of P32 radiolabelled probes	30
Radioactive labelling of oligonucleotides	30
Radiative labelling of DNA fragments	30
Spin column preparation and use	30
2.2.4 Screening of bacterial clones	30
2.2.5 Southern Blots	31
2.2.6 DNA sequencing	32
Preparation of acrylamide gradient sequencing gels	32
2.2.7 Polymerase chain reaction (PCR)	33
PCR screening of recombinant clones	33
Purification of synthetic oligonucleotides	34
2.2.8 Construction of plasmids	34
L κ as basis for further constructs	34
2.2.8.1 The L κ -CDR1 mutated constructs	37
2.2.8.2 The κ 3' enhancer dissection constructs	38
L $\kappa\Delta$ BA[3'E]	38
L $\kappa\Delta$ [F13'EC] and L $\kappa\Delta$ BA[F13'EC]	38
L $\kappa\Delta$ BA[3'EC]	38
2.2.8.3 The κ intron enhancer deletion construct (L $\kappa\Delta$ BA[Ei])	39

2.2.8.4 The MAR deletion construct (LκΔBA[MAR])	39
2.2.9 Cell culture techniques	40
Freezing and thawing cell lines	40
2.2.10 Isolation and analysis of Peyer's patches cells	40
Isolation of cells	40
Fluorescence activated cell sorting	40
2.2.11 Transfection of mammalian cells and analysis of gene expression.....	40
Generation of stable transfectants	40
Ribonuclease protection assay	41
Northern blot analysis	42
2.2.12 Enzyme-linked immunoabsorbent assay	42
Inhibition assays for detection of transgenic κ expression.....	42
Inhibition assays to quantitate mouse κ expression in supernatant of stable transfected S107 cells.....	43
2.2.13 Transgenic mice and hybridomas	43
Preparation of the DNA fragments for microinjection	43
Immunisation protocols	44
2.2.14 Computer analysis	44

Chapter Three

Transgenes

3.1 Introduction	45
3.2 Transgenic lines to study the role of the DNA sequence and secondary structure in hypermutation: Lκ(WT&7M), LκNG and Lκ7M/TCA&Δ[Li].....	47
3.2.1 Lκ(WT&7M) line	47
3.2.2 LκNG line.....	48
3.2.3 Lκ7M/TCA&Δ[Li] line	50
3.3 Transgenic lines to study the role of the 3' enhancer core and its flanking sequences: LκΔ[3'EC] and LκΔ[Fl3'EC]	52
3.3.1 LκΔ[Fl3'EC] line	52
3.3.2 LκΔ[3'EC] line	54
3.4 Transgenic lines to study the role of the DNA elements within the J-C intron: LκΔ[MAR] and LκΔ[Ei]	56
3.4.1 LκΔ[MAR] line	56
3.4.2 LκΔ[Ei] line.....	58
3.5. Expression analysis	58
3.5.1 Transgene κ chain expression in hybridomas derived from the transgenic lines	59
3.5.2 Transgene κ chain expression in the serum of the transgenic mice	59
3.5.3 Transgene κ chain expression in stably transfected NS0 cells.....	64
3.5.4 Transgene κ chain expression in stably transfected S107 cells.....	66
3.6 Discussion	68

Chapter Four

Influence of DNA structure and primary sequence in somatic hypermutation

4.1 Introduction	72
4.1.1 Modifications introduced in the CDR1 of the V κ transgenes	74
4.2 The κ 3' enhancer drives hypermutation from a position 5' of the target area	76
4.3 The modifications in the DNA sequence and secondary structure did not affect the rate or the nucleotide substitution preferences of hypermutation	78
4.4 The DNA sequence has a major influence on hotspots	82
4.4.1 A new intrinsic hotspot in Ser31	84
4.4.2 Ser 31 as TCA is not a hotspot	87
4.4.3 'TA' motif as mutational target	89
4.5 The secondary structure has a lesser effect in the definition on hotspots	91
4.6 DNA sequence modifications and affinity maturation	92
4.7 Discussion	94

Chapter Five

Controlling elements in the κ locus

5.1 Controlling elements in the J-C κ intron	96
5.1.1 Hypermutation of Δ Ei κ transgenes	100
5.1.2 Hypermutation of Δ MAR κ transgenes	107
5.2 Controlling elements in the κ 3' enhancer	111
5.2.1 Hypermutation of L $\kappa\Delta$ [Fl3'EC] transgenes	112
5.2.2 Hypermutation of L $\kappa\Delta$ [3'EC] transgenes	114
5.3 Discussion	117

Chapter Six

Discussion

6.1 The core of the κ 3' enhancer is able to drive expression and hypermutation of the κ light chain but its flanking regions are not essential for either of these processes	119
6.2 The deletion of the κ intron enhancer did not abolish expression or hypermutation	120
6.3 The MAR is essential for hypermutation	120

6.4 The role of the primary DNA sequence and hairpin loops on hotspots	121
6.4.1 The primary sequence is important for the specific location of hotspots.....	122
6.4.2 The secondary structure has a lesser contribution on hotspots	122
6.5 The downstream position of the 3' enhancer with respect to the target area is not essentialfor hypermutation	122
6.6 Not all the transgene copies accumulate the same amount and type of mutations	122
6.7 The mechanism of somatic mutation.....	123
6.8 Conclusions and outlook	124
References Cited	126
Appendices	152

1.1 General background

1.1.1 The beginning of modern immunology

The concept of "protective immunity" induced by prior contact with disease was documented as early as 430 BC by the Greek philosopher Thucydides, who chronicled the plague of Athens, noting that '...no one caught the disease twice or, if he did, the second attack was never fatal'. This understanding, based on pure observation, was independently put into practice at the same time in other parts of the world. Indeed, the ancient Africans, Chinese and Turks practised different types of smallpox inoculation to protect from a recurrence of the disease. A more rational fight against infectious diseases began in the late 18th century when Edward Jenner uncovered the principles of active immunisation and introduced vaccination against smallpox (Jenner, 1789). In 1900, Ehrlich proposed that the binding of foreign toxins to side chains on the surface of cells stimulates the cells to produce serum antitoxins (Ehrlich, 1900).

1.1.2 The instructive and selective theories of antibody formation

In the 1930's, Ehrlich's theory was challenged by Landsteiner's demonstration that antibodies bind an almost infinite number of natural as well as artificial antigens (Landsteiner, 1945). The requirement of an enormous repertoire of pre-existing antibodies to recognise the huge variety of natural and artificial antigens rendered Ehrlich's proposal implausible. Instead, it was suggested that the specific binding sites of the antibodies gain their three dimensional configuration under the direct influence of the antigenic determinant (Mudd, 1932; Pauling, 1940). However, this 'instructive theory' failed to explain processes such as memory and tolerance.

In 1957, Burnet proposed the clonal selection theory according to which cells producing antigen specific antibody are selectively stimulated to proliferate and produce antibodies. This theory proposes that the specific antibodies are expressed prior to exposure to the antigen. The antigen then selects a cell expressing a single specificity out of a pool of cells committed to different specificities. All the cells of the resulting clone produce antibodies with the same specificity.

1.1.3 The structure of immunoglobulins

Early work, based on physicochemical techniques, localized antibody activity to two classes of proteins in the γ globulin fraction of serum (Heidelberger and Pederson, 1937; Tiselius and Kabat, 1939). The major component of antibody activity localized to a subfraction with a sedimentation coefficient of 7S whereas the second component localized to the 19S subfraction. Subsequently, a third protein, γ 1A was detected by immunoelectrophoresis (Grabar et al., 1956). These proteins were named immunoglobulins (Heremans, 1960) and the 7S, 19S and γ 1A fractions were named IgG, IgM and IgA classes respectively (World Health Organisation, 1964). Later two additional classes of immunoglobulins, IgD and IgE were identified.

Further understanding of the immunoglobulin structure resulted from the digestion of rabbit γ globulin with the proteolytic enzyme papain and revealed that the antibody molecule can be split into two identical antigen binding fragments (Fab) and another fragment which easily crystallises (Fc) (Porter, 1959). Denaturation of monoclonal material (purified from the serum of patients with myelomas) with urea revealed that immunoglobulin molecules consist of four separate protein chains. Two identical heavy chains (approximately 440 amino acids long) and two identical light chains (approximately 220 amino acids long) are connected by disulphide bonds (Edelman and Poulik, 1961). On the basis of serological characteristics the light chains could be divided into two types: κ and λ (Nisonoff et al., 1975), the majority of light chains in the peripheral blood being of the κ type.

The variability in the binding of antibody to haptens of different sizes suggested a heterogeneity of antibody combining sites analogous to the heterogeneity of binding properties (Schlossman and Kabat, 1962). A key proof of the heterogeneity of immunoglobulins came from amino acid sequence data. Many patients with myeloma excrete Bence-Jones protein (free immunoglobulin light chain) in their urine. The comparison of a large number of those κ light chain sequences showed very little variation between the carboxyl terminal constant (C) regions, whereas the amino terminal variable (V) region showed high variability (Hood and Talmadge, 1970). When comparing the degree of amino acid variability within the V regions the greatest amount of variability was evident within the three regions known as complementarity determining regions (CDR1, CDR2 and CDR3). These CDRs are flanked by less variable framework regions (Wu and Kabat, 1970).

The polypeptide chains of immunoglobulins are composed of repeated homology units (Hill et al., 1966; Edelman and Gall, 1969). Each of the units consist of two stacked anti parallel β sheets surrounding an internal space filled with hydrophobic amino acid side chains (Amzel and Poljak, 1979). Cysteine residues at the flanks of the units form intra-

chain disulphide bonds stabilising the tertiary structures which commonly referred to as immunoglobulin folds. Electron microscopic studies of immunoglobulin-antigen complexes made it possible to deduce the structure of antibody molecules. The polypeptide chains are arranged in a Y shape with the antigen binding site made up by both heavy and light chain on the amino terminal side of the Fab fragment (Valentine and Green, 1967).

1.2 Affinity maturation

After immunisation, especially after repeated immunisations, the affinity of the antibodies in the serum increases dramatically with time. This increase in the quality of serum antibody, which focuses the immune response to a particular antigen, is called affinity maturation.

One of the first observations regarding affinity maturation was that the anti-diphtheria toxin antibodies in serum progressively increase in avidity over time (Jerne, 1951). Several groups reported an increase in affinity over a period of time (Eisen and Siskind, 1964; Klinman et al., 1966) and also a dependency of the maturation on the original dose of antigen (Eisen and Siskind, 1964; Goidl et al., 1968).

The next major step in the understanding of affinity maturation was to link the antibodies bearing the affinity to the cells which produce them. Anti-DNP antibodies synthesised by a suspension of lymph node cells, which had been obtained at various intervals from immunised rabbits, progressively increased in their affinity (Steiner and Eisen, 1967a, b). This led to the suggestion that the changes can be understood as the result of a selection by the antigen, on the basis of its interaction with pre-existing antibody molecules of those cells synthesising the highest affinity antibody. Low affinity cells failing to capture the antigen are not stimulated to proliferate. Consequently they disappear from the antibody forming cell population (Siskind and Benacerraf, 1969). This suggestion already captured the basis of our understanding of affinity maturation today (see Section 1.5 'Cellular dynamics of affinity maturation'): the preferential selection of those B cells which produce antibodies with the best binding characteristics.

1.3 Origins of immunoglobulin diversity

1.3.1 Germline versus somatic theory

Two fundamentally different theories, both attempting to explain the genetic mechanism to achieve the high variability of immunoglobulins, were suggested in the 1960's and 1970's. The germline theory postulated that all cells have the same range of genes, a choice of which is expressed in each cell. Consequently every cell must contain all the

genes encoding all the necessary antibodies. The somatic theory rejected this assumption, because it would require millions of genes. Instead it postulated that the immunoglobulin genes evolve by mutation from relatively few germline genes during somatic development.

In 1965 Dreyer and Bennett (1965) made a radically different suggestion. They proposed that the genetic information for the variable and constant regions are encoded by two separate unconnected gene segments. Further they speculated that there are hundreds of variable region genes but only a few genes for the constant regions. According to this theory one of the variable genes was brought together with one of the constant region genes by a process of somatic rearrangement.

1.3.2 Recombination of immunoglobulin genes

The first indications of somatic recombination came from Southern blots in which mouse myeloma λ light chain cDNA was hybridised with cloned genomic DNA. The result showed that the λ light chain gene segments were separated in the germline, whilst being contiguous on mRNA level (Hozumi and Tonegawa, 1976). Sequence analysis of the V (variable) gene segment revealed that a little piece was missing, which led to the identification of another group of J (joining) gene segments (Brack et al., 1978). From this it was concluded that each λ light chain is assembled by a combination of V, J and C gene segments. Shortly afterwards, the κ light chain and the heavy chain were analysed analogously (Seidman et al., 1979; Early et al., 1980; Sakano et al., 1980).

A multi step process yields the contiguous information for a functional immunoglobulin chain. Initially single V, (D) and J gene segments are brought together. Then the DNA is transcribed into RNA which contains a VJ (or VDJ) complex, the C gene segment and the intron. Finally, the intron is cut out by RNA splicing and the resulting mRNA is translated into protein.

All the recombinationally active gene segments are flanked by conserved recombination signal sequences that consist of a heptamer and an AT-rich nonamer separated by a spacer of either 12bp or 23bp (the equivalent of one or two turns of the DNA helix). The presence of a pair of recombination signals is all that is required to render a DNA molecule a substrate for V(D)J recombination (12/23 rule) (reviewed by Tonegawa, 1983). Together, two lymphoid-specific genes, RAG1 and RAG2 (Schatz et al., 1989; Oettinger et al., 1990) initiate the recombination process by recognising the recombination signal and cleaving between the signal and the coding border. The joining mechanism is imprecise and this innate imprecision generates significant diversity in immunoglobulin polypeptides (Kurosawa and Tonegawa, 1982; Alt and Baltimore 1982).

1.3.3 Diversification of immunoglobulin genes during B cell development

The four major sources of immunoglobulin diversity are: i) combinatorial diversity by rearrangement of multiple germline segments to build a functional unit, ii) imprecise joining during recombination leading to the addition or loss of nucleotides, iii) association of heavy and light chain to form the antigen binding site and iv) somatic hypermutation. The first three diversification steps take place early on in the development of B cells and are independent of stimulation by antigen. The fourth step of diversification - somatic hypermutation - happens after stimulation by the antigen and is thought to be one of the driving forces behind affinity maturation.

1.4. Somatic hypermutation of immunoglobulin genes

1.4.1 First indications of somatic hypermutation

In 1970 Weigert (1970) analysed the amino acid sequence of the $\lambda 1$ light chain of 18 murine myelomas of mice belonging to the same inbred line. Twelve of them had identical sequences. The other six carried different amino acids at certain residues (Weigert et al., 1970). This was the first indication of a mechanism which increases the variability of immunoglobulin genes by somatic hypermutation. With the cloning of the $\lambda 1$ germline gene it was confirmed that the changes were indeed due to somatic mutation rather than gene polymorphism (Brack and Tonegawa, 1977).

The first data on somatic hypermutation came from comparing the amino acid (and later nucleic acid) sequences of immunoglobulins produced by malignant plasma cells with their germline counterparts (Gearhart et al., 1981; Crews et al., 1981).

Although these studies already demonstrated that somatic hypermutation plays a major role in the diversification of antibodies, they did not provide any information about the development of antibody genes during an immune response. The hybridoma technology (Köhler and Milstein, 1975) combined with fast DNA sequencing (Hamlyn et al., 1981) finally made it possible to study the genetic events that operate in response to an antigenic challenge. This approach was exploited to study various other responses, including those to a variety of haptens such as 2-phenyl-5-oxazolone (Kaartinen et al.; 1983; Griffiths et al.; 1984; Berek et al., 1985), (4-hydroxy-3-nitrophenyl)-acetyl (Bothwell et al.; 1981; Cumano and Rajewsky, 1986; Blier and Bothwell; 1987) p-azophenylarsonate (Wysocki et al., 1986) and protein antigens such as influenza hemagglutinin (McKean et al., 1984)

1.4.2 The rate of somatic hypermutation

The rate of somatic hypermutation is estimated as approximately $10^{-3}/\text{bp}/\text{generation}$ (McKean et al., 1984; Sablitzky et al., 1985). This is approximately ten thousand fold higher than the rate of spontaneous mutation in higher eukaryotes (Adetugbo et al., 1977; McKean et al., 1984). This mutation rate seems to allow an adequate accumulation of mutations. A higher rate of mutation might result in the loss of clones due to deleterious mutations, a lower rate would be less efficient in obtaining the desired mutations (Allen et al., 1987). Therefore it is not surprising that the estimate remains almost the same even if the calculation only accounts for silent changes (Berek and Milstein, 1988) or has been obtained at different stages of the primary and secondary response (Rada et al., 1991; Weiss et al., 1992).

1.4.3 Target area of somatic hypermutation

Somatic hypermutation of Ig genes is a tightly controlled, site directed process. The nucleotide changes are targeted to the region surrounding the rearranged V gene segment of heavy and light chain loci (Gorski et al., 1983). Mutations do not extend significantly upstream beyond the promoter or significantly downstream into the constant region exons (Gearhardt and Bogenhagen, 1983; Lebeque and Gearhart, 1990; Both et al., 1990; Weber et al., 1991; Rothenfluh et al., 1993).

Occasional mutations have been described in a mouse C λ exon, perhaps because in the mouse λ locus the C region is closer to the joining (J) cluster than in other immunoglobulin loci (Motoyama et al., 1991). More recently the 5' border has been located in the intron between the leader peptide and the variable region in κ chains (Rada et al., 1994; Rogerson, 1994). A pyrimidine (mostly T) rich motif occurs at the boundary position, but its significance is unknown.

1.4.4 The timing of somatic hypermutation

Somatic hypermutation is switched on very shortly after primary immunisation. The first mutations can be detected 4 days after local immunisation in cells of the draining lymph nodes (Leanderson et al., 1992) and six to seven days after immunisation in the cells of the spleen (McHeyzer et al., 1991). The cells appear to accumulate mutations for up to 21 days after primary immunisation (Rada et al., 1991; Weiss et al., 1992). Diversification by somatic hypermutation occurs in a step-like manner (Brown et al., 1992), suggesting that several rounds of mutation occur. Secondary immunisation leads to the expansion of the memory compartment and seems to reactivate the hypermutation mechanism (Berek and Milstein, 1988; Blier and Bothwell, 1988; Rada et al., 1991).

1.4.5 Intrinsic features of somatic hypermutation

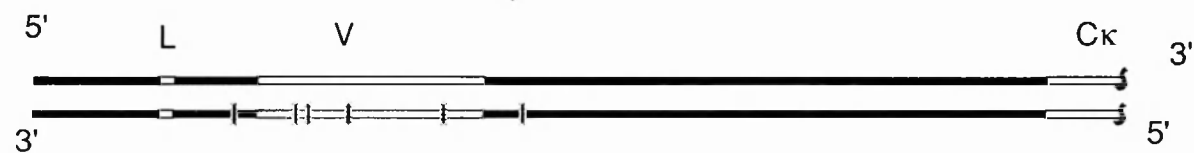
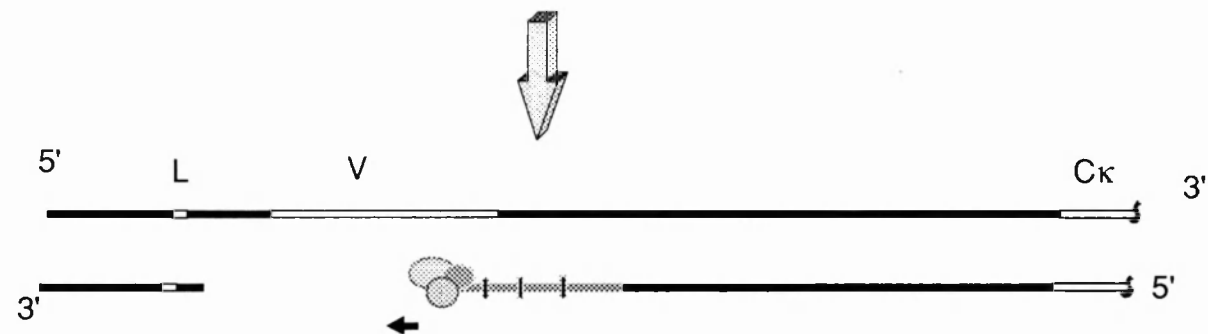
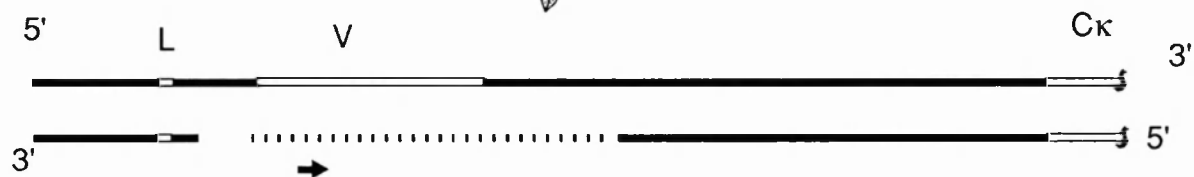
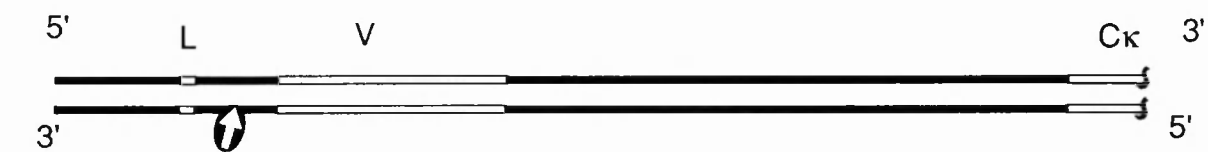
Large variations in the mutation (basically nucleotide substitutions) frequency at which mutants are found within the targeted area. Certain positions form favoured sites of mutations (hotspots) and others are rarely mutated (coldspots). Although large databases are only available for fewer examples, mutational hotspots were found in all cases even when the V region of the transgene was substituted by heterologous sequences (Yélamos et al., 1995). In particular the CDR1 is usually more strongly mutated than other regions (Malipiero et al., 1987; Betz et al., 1993b), maybe because it contains a large accumulation of mutational hotspots. However, hotspot clusters were also found in the joining/constant region intron as well as in the CDR2, CDR3 of the $\lambda 1$ gene (González-Fernández et al., 1994b).

The mutations are also not generated randomly in terms of the base substitution preference. Instead they exhibit a pronounced bias, with purine residues on the coding strand being preferred over pyrimidines as target for mutations. This "strand polarity" can not be attributed to an uneven base composition of the targeted area. In addition, transversions occur at lower frequency than transitions. The directionality of mutation often allows the discrimination between antigen-selected or intrinsic hotspots (Betz et al., 1993a).

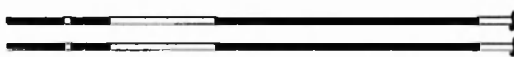
A striking feature of many coding for serine major intrinsic hotspots is that they constituted AGC or AGT triplet (Betz et al., 1993a). A preference for the usage of AGY codons in the CDRs (particularly in the CDR1) and for TCNs in the frameworks has been observed in human V genes (Wagner et al., 1995). Codon bias has also been observed in *Xenopus* and sheep (Schwager et al., 1989; Reynaud et al., 1995). It suggests that the DNA sequence of germline V genes has evolved in response to selection for appropriately targeted mutability.

The hotspots existence has been attributed to peculiarities of DNA structure, including palindromes, repeats (Milstein et al., 1986; Golding et al., 1987; Kolchanov et al., 1987) and primary sequence motifs. The latter category includes the motifs TAA, PuGPyA/T (Rogozin and Kolchanov, 1992), CAGCT/A and AAGTT (Betz et al., 1993a). The hotspots found in the non-immunoglobulin sequences also fall within the intrinsic hypermutation hotspot consensus (Yélamos et al., 1995).

Mutational hotspots are only in part created by their sequence since they are not found in all the positions conforming to the hotspot consensus. The hotspots found in the non-immunoglobulin sequences are not located at a conserved distance from the transcription start site (Yélamos et al., 1995) and may therefore be a consequence of neighbouring DNA sequence, or structure, such as palindromes (Milstein et al., 1986; González-Fernández et al., 1994b).



REPLICATION



1.4.6 Sequences needed to recruit hypermutation

Transgenic mouse experiments have revealed most of what is known about the *cis*-acting DNA elements required to recruit hypermutation. The rate of hypermutation was hardly affected when the V κ promoter was substituted with the promoter of the β -globin gene (Betz et al., 1994). The V region was replaced by non-immunoglobulin sequences and hypermutation was not impeded (Yélamos et al., 1995). A big part of the sequence between the C κ and the 3' enhancer as well as a small fragment in the J-C intron were removed and hypermutation was not affected (Betz et al., 1994; Yélamos et al., 1995). On the other hand, removal of either the κ 3' enhancer or the κ intron enhancer/MAR caused a drastic reduction in hypermutation, which is particularly marked in the latter case (Betz et al., 1994). Whether the effect of the enhancers in hypermutation is secondary to their activity as transcription activator, or whether they stimulate hypermutation through an independent (but parallel) process is a fundamental question that need to be resolved to understand the mechanism of somatic mutation.

1.4.7 Mechanism of somatic hypermutation

The molecular details of how somatic hypermutation works are not yet known. Several models have been proposed on the theme of an error-prone repair system (Brenner and Milstein, 1966) or misaligned templates (Malipiero et al., 1987). A version of the first model (Fig. 1.1) involves a singlestrand nick in the transcribed strand, in a site upstream of the V region (which determines the 5' boundary of hypermutation) followed by exonuclease digestion and replacement by an error-prone DNA polymerase such as DNA polymerase β (Kunkel and Alexander, 1986).

Figure 1.1 Model of somatic hypermutation in κ light chain (opposite page).

A version of the error-prone repair model proposed by Brenner and Milstein, 1966 (taken from Milstein and Rada, 1995). A specific single-strand nick is introduced in the transcribed strand. This is followed by exonucleolytic trimming of the nicked strand with subsequent error-prone repair. Mutations are introduced in one strand and are fixed during DNA replication. This accounts for the observed strand polarity, which is deduced from the biases intrinsic to hypermutation. In this version of the model the preferred cleavage site is located upstream of the V region. The exonucleolytic activity 'falls off' after an average of about 500-1000bp. Alternatively there could be a large number of cleavage sites in the J-C intron of the non transcribed strand and a sharp stop of the nucleolytic cleavage in the leader intron.

This model is compatible with the strand polarity, the sharp boundary at the 5' end and the decay at the 3' end. It also finds considerable support in the emerging similarities between hypermutation and excision repair of damaged DNA, e.g. the preferential repair of the transcribed DNA strand is a recognised feature of nucleotide excision repair (Hanawalt and Mellon, 1993).

Some variations of the model include an RNA intermediate and reverse transcription and insertion (Steele and Pollard, 1987) or templated gene conversion (Maizels, 1989). While being the main mechanism of somatic diversification in the chicken (reviewed in McCormack et al., 1991) the gene conversion mechanism is unlikely at least in mice and man (Gearhart and Bogenhagen, 1983; González-Fernández and Milstein, 1993; Milstein, et al., 1992; Wysocki et al., 1990). Thus in these species mutations extend the flanking regions 3' of the V, where no donor homologous sequences are present in the genome, and no concordant pattern of somatic mutants with genomic diversity was found. An additional evidence that ought to put to rest this possibility is the demonstration that several non immunoglobulin sequences (which are not present in the mouse) replacing the V segment hypermutate (Yélamos et al., 1995).

A model that links DNA replication with a process that targets somatic mutation has also been proposed (Rogerson et al., 1991). However, there is no evidence for the specifications postulated by this model and it has been unsubstantiated by the finding that the somatic hypermutation was affected when the transcriptional enhancers and MAR were removed, whereas replication of the transgenes continued (Betz et al., 1994). Somatic mutation has been described also in sheep, where it serves to diversify the preimmune repertoire (Reynaud et al., 1991) similar to the diversification by gene conversion that occurs in the Bursa of chicken. Taken together these observations indicate that a similar molecular mechanism based on the same basic DNA repair machinery might be used at different times during B cell development by different species for different purposes.

1.5 Cellular dynamics in affinity maturation

1.5.1 The primary response and the germinal centre formation

In the immune response against T-cell-dependent antigens the first detectable (48 hours after immunisation) expansion of antigen-specific B cells takes place in the periarteriolar lymphoid sheaths (Gray et al., 1986; Jacob et al., 1991). These are areas in peripheral lymphoid organs (the peripheral lymphatic tissues and the spleen) where T cells accumulate and are organised around specialized endothelial vessels. During this early expansion some B cells differentiate into plasma cells and start producing low affinity IgM antibodies.

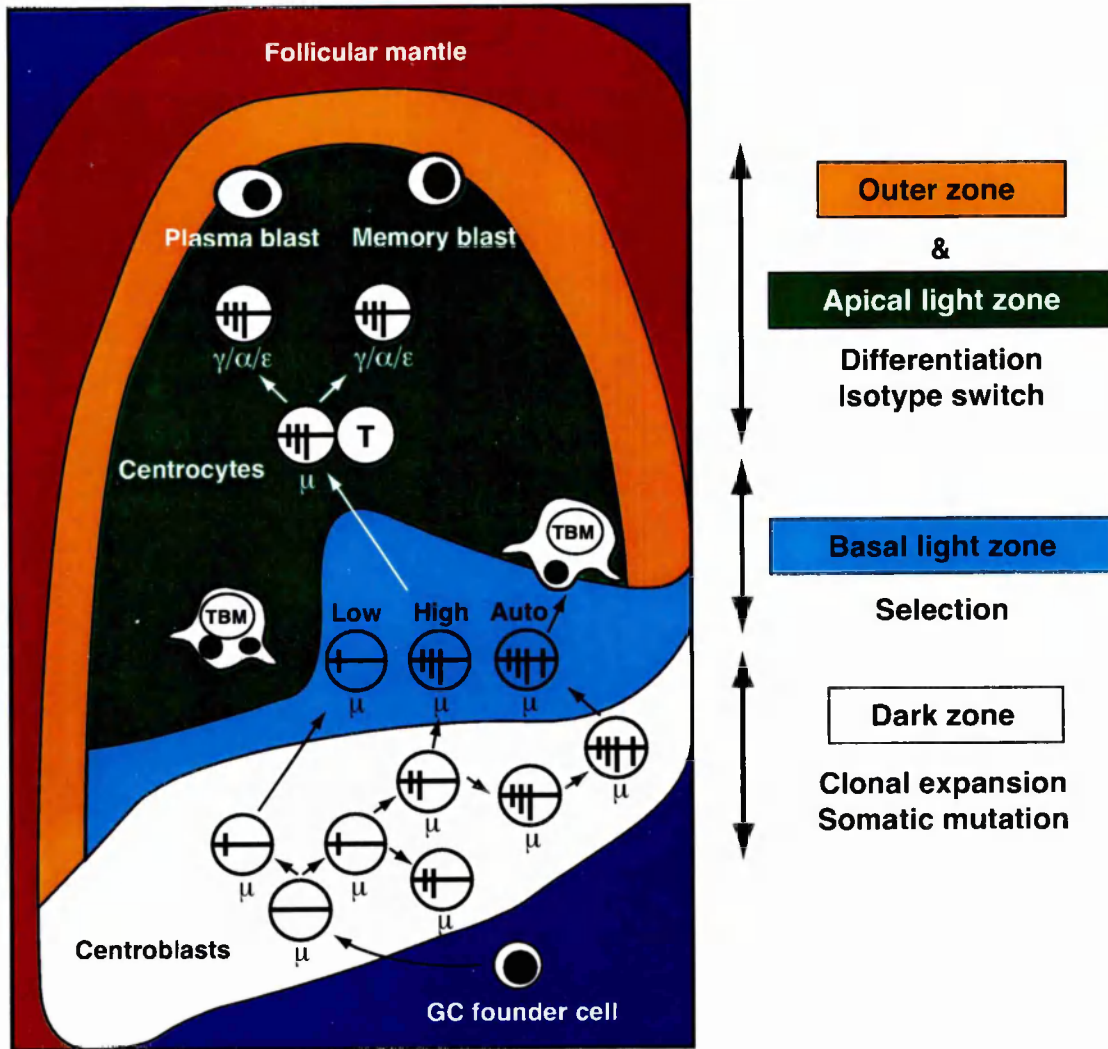


Figure 1.2 Functional compartments in germinal centres (taken from Liu et al., 1996).

A few days after immunisation B cells begin to migrate into the primary follicles of the peripheral lymphoid organs (Kroese et al., 1990) where they undergo extensive proliferation and form germinal centres (GC), which appear towards the end of the first week after challenge and provide the location for hypermutation, antigen selection and isotype switching (MacLennan et al., 1992; Liu et al., 1992; Liu et al., 1996).

Germinal centres are highly organised structures that develop around the follicular dendritic cell (FDC) network. Their formation depends on the presence of Ag-Ab complexes trapped in the interdigitating processes of FDCs (Nossal et al., 1968; Tew and Mandel 1978; Klaus et al., 1980). Several cellular compartments have been identified histologically (Hardie et al., 1993) and a number of studies that suggest functional differences between them.

1.5.2 Germinal centre - microenvironment for hypermutation, antigen selection and isotype switching

A simple model for the maturation pathway of a IgM-expressing B cell in the germinal centre is shown in Fig 1.2. An IgM⁺ B cell migrates to the dark areas of the GC, where they expand exponentially. These cells later become centroblasts, which express little or no surface Ig. Centroblasts are thought to be the cells in which the somatic mutation occurs (MacLennan et al., 1992; Küppers et al., 1993; Pascual et al., 1994). They transform into nonproliferating centrocytes, which re-express antigen receptor on the surface albeit at low levels, migrate towards the light zone of the GC and undergo apoptosis unless further stimulated by additional signals (Liu et al., 1989). The Ag-Ab immune complexes on the FDC are the main signal that prevents apoptosis and selects high affinity mutants specific for the antigen. Additional signals as T-cell derived (IL-2, IL4 and IL10) cytokines and those delivered through the interaction of CD40 on B cells with its ligand (CD40L) also are involved in preventing apoptosis.

While the low affinity and autoreactive mutants die by apoptosis, high affinity mutants pick up the antigen and process it on their migration pathway to the apical light zone and the outer zone. In these areas the selected centrocytes encounter and present antigen to antigen-specific T cells. T cells are induced to express CD40 ligand (Lederman et al., 1992; Casamayor-Palleja et al., 1995) and secrete cytokines which are both key elements for the induction of isotype switching. This cognate T-B interaction results in the expansion and the isotype switch of high affinity centrocytes.

Finally, they differentiate into memory B cells in the presence of prolonged CD40 ligand signalling and into plasma cells when CD40 ligand signalling is removed (Liu et al., 1992; MacLennan et al., 1994; Arpin et al., 1995).

Antigen selection seems to operate from the very beginning of germinal centre formation. Somatic mutation, although detectable at day six, peaks somewhat later (well into the second week after challenge) and antibodies derived in early stages of primary

responses reflect proliferation without hypermutation (Griffiths et al., 1984; Källberg et al., 1993; Jacob et al., 1993).

Isotype switching seems to initiate after B cells have undergone somatic mutation and positive selection. This allows the acquisition of a distinct Ig isotype, which confers effector functions, while keeping the high affinity for the selecting antigen. A small number of the cells may undergo isotype switching without somatic hypermutation and vice versa (Kaartinen et al., 1983; Cumano and Rajewsky, 1985; Jacob and Kelsoe, 1992; Pascual et al., 1994). This indicates that they are two independent processes.

1.5.3 B cell memory

Memory B cells can be defined as surface Ig expressing B lymphocytes which have been selected in a pathway of antigen-driven proliferation and somatic hypermutation. They also appear to have already switched their isotype (Okumara et al., 1976; Coffman and Cohn, 1977). On renewed antigenic challenge they produce the secondary immune response (Kocks and Rajewsky, 1989). Whilst the virgin B cells are characterised by the expression of surface IgM (sIgM) and sIgD, the B cells involved in the secondary response usually express antibodies of the IgG, IgA or IgE class. There is increasing evidence supporting the hypothesis that memory cells recirculate through GCs to generate even higher affinity clones and to switch to downstream isotypes (Berek and Milstein, 1987; Kepler and Perelson, 1993; Pascual et al., 1994; Feuillard et al., 1995).

It is still unclear whether the persistence of the B cell memory requires ongoing stimulation and selection by the antigen, or if it simply reflects the generation of stable long-lived memory cells (for review see Freitas et al., 1986; MacLennan and Gray, 1986). It has been suggested that the *bcl2* gene expression possibly plays an important role in the longevity of memory cells (Vongsakul, 1996).

1.6 The immune response against 2-phenyl-oxazol-5-one

1.6.1 A model for somatic hypermutation

The primary response against the hapten 2-phenyl-oxazol-5-one (phOx) in many mouse strains is dominated by one major idiotype (Mäkelä et al., 1977). Most of the phOx-specific B cells utilise a particular combination of heavy and light chain, termed V κ Ox1 and VHOx1 respectively (Kaartinen et al., 1983; Even et al., 1985). This conservatism in respect of immunoglobulin gene usage renders it an ideal model system to study the development of immunoglobulin genes during an immune response. Somatic mutants with increased affinity to phOx can be detected (Griffith et al., 1984) from early on in the response.

1.6.2 Affinity maturation by somatic hypermutation

Early in the response, most of the V κ Ox1 and VHOx1 genes are expressed in their unmutated germline form. As the response progresses the V regions start to accumulate mutations. Whilst many mutations are specific for the affected immunoglobulin gene, common mutations such as substitutions in His34 and Tyr36 can be observed in most of the antibodies of a secondary or tertiary immune response to phOx. Substitution of Tyr36 to Phe alone does not seem to greatly affect the affinity, but a combination of this substitution with the substitution of His34 by Gln or Asn results in a ten-fold increase in affinity (Berek and Milstein, 1987). The determination of the three dimensional structure of the Fab fragment (Alzari et al., 1990) of an anti-phOx monoclonal antibody (NQ10/12.5) revealed that the residues 34 and 36 of the V κ Ox1 light chain directly interact with the hapten. It is therefore not surprising that mutations in these key contact residues alter the affinity for phOx. In contrast, mutations in many of the other residues have little effect on hapten binding (Alzari et al., 1990). Together with the silent changes, they represent a neutral background resulting from the nature of the mutational process. It has been estimated that only one out of four point mutations has actually been selected by antigen (Berek and Milstein, 1988).

1.6.3 Affinity maturation by repertoire shift

Initially, the response against phOx is dominated by the V κ Ox1/VHOx1 combination, probably as a result of its relatively high incidence in the preimmune repertoire (Milstein et al., 1992). At later stages of the primary response and particularly after a second challenge by the antigen, other germline gene combinations emerge. The V κ Ox1 light chain can frequently be found in combination with various different heavy chains (Berek et al., 1985) such as members of group 5 - VHM21 like or group 1 - VHJ558 like (Dildrop, 1984). The association of the V κ 45 light chain with the VH11 heavy chain is another typical combination found in the secondary response (Kaartinen et al., 1988). The antibodies utilising these alternative combinations often do not have a higher affinity than the mutated V κ Ox1 based antibodies. Instead, differences in the kinetic properties of the binding of the antigen appear to account for their emergence (Foote and Milstein, 1991).

1.7 Regulation of the immunoglobulin gene transcription

The regulated transcription of immunoglobulin genes depends, as in any cell system, on contributions from multiple *cis*-control regions and multiple regulatory proteins. The analysis of the Ig transcription is complicated by the gene rearrangements in immature cells and by the isotype switching events in mature cells. The precise regulatory mechanisms have not been yet clearly established.

1.7.1 *Cis*-acting sequences: promoter and enhancers

The promoter

The immunoglobulin promoters are only active in B-lineage cells. All heavy and light promoters are composed by two main elements, the octamer and the A/T rich sequence (TATATAA) that function as a TATA box (Parslow et al., 1984).

The TATA motif, located 25-30bp upstream of the transcription start site sequence, directs transcription initiation. This activity result from the binding of TATA binding protein (TBP) which is a subunit of a multiprotein complex commonly called transcription factor IID (TFIID).

The "octamer" element consists of nine conserved bases: ATTTGCATNA, with N being a variable base at position 9 (Wirth et al., 1987). It is located approximately 70bp 5' of the cap site in the light chain and in the same position but in an inverted orientation in the heavy chain gene. Synthetic promoters containing an octamer site and a TATA box can direct accurate an B-lineage-restricted transcription (Dreyfus et al., 1987; Wirth et al., 1987). However, the same octamer element drives the expression of a number of ubiquitously expressed genes, including the histone H2B gene (LaBella et al., 1988).

Two POU domain proteins, the ubiquitous Oct-1 and Oct-2 (expressed primarily in B cells), activate transcription by binding the immunoglobulin octamer element (Staudt and Lenardo, 1991). In addition, a B cell restricted coactivator (OCA-B) provides the B cell specificity attributed to the octamer element in the immunoglobulin promoter (Pierani et al., 1990; Luo and Roeder, 1995).

The intron enhancer

The first transcription enhancer elements identified to regulate immunoglobulin gene expression are located in the major intron of the light and heavy chain loci. When put under the control of these enhancers the expression of heterologous genes became lymphoid specific (Banerji et al., 1983; Gillies et al., 1983; Neuberger, 1983; Picard and Schaffner, 1984a; Queen and Stafford, 1984; Reik et al., 1987).

The κ intron enhancer (κ Ei) is inactive in the early stages of B cell development and becomes active during the transition from pre-B to mature B cell (Sen and Baltimore, 1986; Attchison and Perry, 1987).

The best studied binding site in the intron enhancer is the 10 bp κ B site. The constitutive binding of NF- κ B to this site in mature B cells and plasma cells is a critical regulatory event in the differentiation from pre-B to mature B cell. Besides κ B there are three μ E2/5 sites that bind to members of the E2A HLH-transcription factor family (Gimble and Max, 1987; Hromas et al., 1988). Mutational studies indicated that those sites are crucial for the activity of this enhancer (Lenardo et al., 1987).

A 200bp fragment upstream of the NF- κ B binding site has^{been} shown to silence gene expression in non-B cell. It may act to prevent the inappropriate activation of the κ gene by NF- κ B induced in other cells types (Pierce et al., 1991). An additional component of negative regulation at this locus can be ascribed to κ NE, a short sequence immediately 5' of the NF- κ B binding site (Saksela and Baltimore, 1993). Cell type specificity of κ NE was found to be determined by a short element, κ BS, immediately upstream of κ NE. Disruption of the κ BS sequence allowed transcriptional repression by κ NE in B cells. Although the κ Ei also shares some regulatory sites with the heavy chain intronic enhancer (H Ei) they do not function in the same way. Thus the H Ei is active throughout B cell development including early stage pro-B and pre-B cells (Gerster et al., 1986). In addition, it contains a promoter for sterile μ transcripts (Su and Kadesch, 1990) which may be required to generate the sterile μ transcripts for VDJ joining (Ferrier et al., 1990; Chen et al., 1993). The H Ei is flanked by two matrix attachment regions and their combination acts as locus control region (see MAR below).

The 3' enhancer

The search for this enhancer in the κ locus was driven by the observation that high levels of κ light chain protein continue to be expressed in the plasmacytoma S107, which lacks the transcription factor NF- κ B (Atchison and Perry, 1987) thought to be an essential requirement for κ expression. This additional enhancer was found 8.5 kb downstream of the constant region exon, is stronger than the κ intron enhancer (Meyer and Neuberger, 1989) and active in mature B cells and plasma cells but not in pre-B cells, T cells or other cell types (Pongubala and Atchison, 1991).

Most of its activity can be localised to a 132 bp core (Meyer et al., 1990; Pongubala and Atchison, 1991). The 3' enhancer core does not contain κ B sites explaining the expression of the κ genes in S107 cells. The involvement of HLH proteins in the regulation of this enhancer was substantiated by repression of the enhancer through the HLH-inhibitory protein Id (Benezra et al., 1990; Pongubala and Atchison, 1991). Transient transfection assays with the core enhancer revealed that it contains at least three functional DNA sequences (PU.1/NF-EM5, E2A and κ E3'-CRE) (Pongubala and Atchison, 1991; Pongubala et al., 1992; Pongubala et al., 1993; Pongubala and Atchison, 1995). The flanking regions of the core enhancer are involved in the developmental control of the 3' enhancer activity (Pongubala and Atchison, 1991; Meyer and Ireland, 1994). In addition to its function providing further activation of the κ chain promoter in late B cell it may also be important for somatic mutation (Betz et al., 1994). Some rearranged mouse heavy chain genes retained full transcriptional activity despite deletion of their intron enhancer (Klein et al., 1984; Wabl and Burrows, 1984; Aguilera et al., 1985). The search for further enhancer elements in the heavy chain locus led to the

identification of the IgH 3' enhancer first in the rat [25kb 3' of C α] (Pettersson et al., 1990) and later in the mouse [12.5kb 3' of C α secreted exon] (Dariavach et al., 1991; Lieberson et al., 1991). A third transcriptional enhancer, the C α 3' enhancer, which is located very close to the C α region, was identified but not studied in detail (Matthias and Baltimore, 1993).

In contrast to the κ 3' enhancer, the H 3' enhancers is considerably weaker than the intron enhancer. In addition to providing activation of VH promoter in late B cells, this enhancer probably plays a role in activating I-region transcripts that are required for class switching (Pettersson et al., 1990; Coffman et al., 1993; Xu et al., 1993).

1.7.2 Regulation of chromatin structure

The control regions described above appear to be principally responsible for the precise expression level and expression pattern observed for Ig genes. The transcription is regulated by a complex array of ubiquitous and cell specific proteins that are capable of binding to promoter and enhancer elements. These proteins interact with each other and, with the help of coactivators, regulate the formation of a pol II-containing preinitiation complex at the transcription start site. But the fact that in eukaryotic cells genes are assembled into chromatin can not be ignored. The Ig locus is, most likely, assembled into an inaccessible chromatin configuration prior to gene activation (Jenuwein et al., 1993; Forrester et al., 1994) possibly in the form of 30nm filament (Ernst and Smale, 1995).

Many studies on the role of chromatin in mammalian gene regulation have focused on specialized control elements like locus control regions, insulators, matrix attachment regions and methylation. However, many of the elements have not been studied in the Ig genes. Elements that may influence the chromatin structure within the Ig locus during B cell development are discussed below.

Methylation

DNA methylation plays a role in the regulation of tissue-specific gene expression, which is reflected in the organization of methyl groups in somatic cells of the organism. DNA sequences are generally methylated in most cells types, but are unmethylated in tissues in which the gene is actively transcribed (Yisraeli and Szyf, 1984). Thus, the observed genome-wide methylation pattern appears to provide a global mechanism for repressing non active genes in all cell types.

The mouse κ gene has a developmental methylation pattern similar to that of other tissue-specific genes. It is highly methylated in the post implantation embryo (Kafri et al., 1992) and remains modified in most somatic cells. It undergoes differentiation-specific demethylation only in mature B lymphocytes (Matther and Perry, 1983; Storb and Arp, 1983; Nelson et al., 1984). A better understanding of the role of DNA

methylation has emerged from studies of the Ig κ gene by Lichtenstein et al. (1994). In their studies, an *in vitro* methylated plasmid containing a rearranged κ gene was demethylated upon transfection into B cell lines but remained methylated following transfection into pre-B and non-lymphoid cells. Lineage specific and cell stage specific demethylation was induced by a compound *cis*-acting element contained within a 1.6kb domain that is made up of the κ intron enhancer core, the MAR and a sequence located downstream to J κ 5. These intronic sequences induced demethylation in the apparent absence of activated transcription, strongly suggesting that demethylation is not simply a result of transcriptional activation.

Matrix attachment regions (MAR)

DNA within interphase nuclei and mitotic chromosomes is organised into topologically constrained looped domains. Certain DNA sequences remain preferentially associated with the matrix fraction, whereas others are released into the soluble supernatant fraction after employing lithium diiodosalicylate and high salt extraction protocols. Matrix or scaffold attachment regions (MARs or SARs) have been mapped using these procedures in and around many different genes in a wide range of organisms (Phi-Van and Stratling, 1988; Gasser and Laemmli, 1986; Cockerill, 1990).

MARs are typically detected in DNA segments that are AT rich. MARs contain short clusters of certain characteristic A+T-rich sequences (AATATATTT and variations) that seem to be required for matrix interaction. MARs also contain sequences matching the consensus for Topoisomerase II cleavage (GTNA/TAC/TATTNATNNA/G). MARs are often localized close to DNA sequences that play a key role in transcriptional regulation (Gasser and Laemmli, 1986; Cockerill and Garrard, 1986; Cockerill et al., 1987). Structural features in MARs include bending (von Kries et al., 1990), a narrow minor groove due to oligo dA tracts (Adachi et al., 1989) and single-strandedness (Probst and Herzog, 1985). Indeed, MARs from different species become stably base-unpaired over an extended length when subjected to the torsional stress of negative supercoiling: the greater the superhelical strain the more unwinding occurs (Kohwi-Shigematsu and Kohwi, 1990; Bode et al., 1992). This inherent unwinding property was shown to be important for binding to the nuclear matrix and for the augmentation of gene expression in stable transformants (Bode et al., 1992). This property may be important in relieving negative superhelical strain that would arise in looped DNA and in altering torsional stress during transcription. Furthermore, a 1.6kb fragment containing the κ E, the MAR, and the 5' flanking sequence of the MAR can induce tissue and cell stage specific demethylation of the rearranged κ light gene (Lichtenstein et al., 1994). The removal of the MAR element had a marked effect on the demethylation activity.

MARs also act as locus control regions (LCR), which effect the accessibility of a genetic locus. The Ig μ LCR consists of a core enhancer which is flanked on either sides by MARs (Jenuwein et al., 1993). The intact μ LCR is able to induce DNase-I-hypersensitive sites and confer accessibility of a linked T7 promoter to the T7 RNA polymerase in isolated nuclei. In contrast the core enhancer on its own is sufficient to confer T7 promoter accessibility, but does not induce formation of a detectable DNase-I-hypersensitive site.

The analysis of the Ig μ LCR was extended to an immunoglobulin transgene (Forrester et al., 1994). Like the globin LCR, the intact μ LCR confers position-independent, cell-type specific expression of the immunoglobulin reporter. However, when the flanking MARs are deleted, the level of tissue-specific expression is substantially reduced in most transgenic lines and the gene exhibits a position dependent variability not evident with the intact μ LCR. Differences are also observed at the level of chromatin organization. The intact μ LCR induces the formation of an extended DNase-I-hypersensitive chromatin domain with hypersensitive sites in the core enhancer and at the μ promoter. When the MARs are deleted, the core enhancer is able to induce only local effects on chromatin structure (the enhancer but not the promoter becomes hypersensitive and no extended DNase-I-sensitive domain is formed). This suggests that MARs are required either to mediate communication between the core enhancer and the μ promoter, or to propagate an enhancer-induced alteration in chromatin structure throughout the entire transgene domain.

The same MARs that flank the μ enhancer core have been also implicated in negative regulation through cell specific anchorage of expressed genes in a region of high transcription factor concentration around the nuclear matrix (Zong and Scheuermann, 1995).

1.7.3 Regulation of the immunoglobulin loci during the B cell development

The expression of immunoglobulin genes is unique to and essentially defines the B cell lineage. The Ig μ locus appears to be activated at a very early stage of B cell development, preceding V(D)J recombination and possibly preceding commitment to the B cell lineage. The initial activation of the locus results in the change in methylation and DNase I hypersensitivity of Ig μ regulatory regions (Mather and Perry, 1983; Storb et al., 1983). Transcripts initiating within the heavy chain locus are detectable at an early progenitor B cell stage (Yancopoulos and Alt, 1985, 1986). Following gene rearrangement, mature transcripts increase throughout the subsequent pre-antigen recognition stages of differentiation (Perry and Kelly, 1979). Ig mRNA is further induced in an antigen and T-cell dependent manner.

Immunoglobulin gene rearrangements occur in an ordered fashion during B cell development. In mice, DJ heavy chain rearrangements in pro-B cells are followed by V

to DJ rearrangement in immature pre-B cells and finally light chain rearrangement in mature pre-B cells (Hardy et al., 1991). The membrane bound form of μ heavy chain is a signal for allelic exclusion - the suppression of further VDJ rearrangements in the heavy chain locus (Nussenzweig et al., 1987; Reth et al., 1987; Manz et al., 1988; Kitamura and Rajewsky, 1992; Papavasiliou et al., 1995).

Usually a B cell expresses a single form of light chain of either the λ or κ type (isotypic exclusion). The fact that most κ expressing B cells retain their λ loci in germline configuration, whilst most λ expressing cells contain two non-productively rearranged κ loci, indicates that rearrangement of the κ loci usually precedes rearrangement at the λ loci (Coleclough et al., 1981; Hieter et al., 1981; Korsmeyer et al., 1981; Lewis et al., 1982).

Isotype switching is another process that occurs in the immunoglobulin loci during the B cell development. It moves the variable heavy gene segment (VDJ) to associate with a different constant region. The expression of sterile transcripts (I γ , I α , I ϵ) is essential for the induction of this process (Coffman et al., 1993; Zhang et al., 1993) and it seems to be directed by cytokines including IL-4 and by engagement of CD40 with its ligand (Kühn et al., 1991; Clark and Ledbetter, 1994; Delphin and Stavnezer, 1995).

1.8 Methodological approaches for analysing somatic hypermutation

1.8.1 The hybridoma approach

The antibody response was initially studied by direct sequencing of mRNA from immunoglobulin genes in hybridomas at different times after immunisation (Griffith et al., 1984; Rudikoff et al., 1984; Sablitzky et al., 1985; Cumano and Rajenwsky, 1986; Malipiero et al., 1987; Manser et al., 1987). Direct mRNA sequencing has been replaced by procedures involving a polymerase chain reaction (PCR) amplification step followed by cloning and sequencing of multiple clones.

This approach still remains the best way to relate the affinity and the amino-acid sequence changes (Berek and Milstein, 1987). In addition, hybridomas provide clear evidence of a common clonal origin of many of the shared mutations found in different monoclonal antibodies. They also allow the analysis of other genes in the same clone, for instance the non-expressed Ig genes (Weiss and Wu, 1987), non allelically excluded transgenes (Sharpe et al., 1991) or multiple copies of identical or mixed transgenes (Rogerson, 1991; Yélamos et al., 1995). As additional sampling can be taken from the same clone, there are no PCR errors in the data bases collected from hybridomas.

This approach, however, is a time and labour consuming process, which generates only small data bases. In addition cells are not immortalised at random (blast germinal centre B cells seem to be immortalised preferentially). The number of mutations detected by hybridomas may therefore be smaller than that present in the memory cell population.

1.8.2 The PCR approach

Amplified gene fragments from populations of B cells can be directly cloned into M13 vectors and sequenced (Orlandi et al., 1989). A major advantage of this approach is its speed and simplicity enabling the generation of a large data bases. In addition it is not restricted to any particular cell subpopulation. It has been used to analyse genes from selected B cell subpopulations (Weiss and Rajewsky, 1990; Rada et al., 1991; Schitteck and Rajewsky, 1992), cells taken from histologically defined areas of germinal centres (Jacob et al., 1993; Ziegner et al., 1994), and also from single B cells or B cell clones (McHeyzer et al., 1991; Küppers et al., 1993). However, the misincorporation errors and artefacts (crossover-like events) produced during the PCR amplification must be carefully considered. Also, the mutation pattern in the heavy and the light chain can not be study in the same clone.

1.8.3 The transgene approach

In the absence of cell lines that reliably reproduce the somatic mutation seen *in vivo*, the use of Ig transgenic mice was introduced as a model to study the intrinsic features as well as the molecular mechanism of this process (O'Brien et al., 1987). κ transgenes are subject to somatic mutations at the same rate and with the same characteristics found in the endogenous counterparts when the relevant flanking regions are included (Sharpe et al., 1991). One major advantage is that there is no question about the origin of the germline gene from which the mutants arise.

To circumvent the labour of sequencing mutated V regions, reporter genes assays were developed to permit rapid identification of hypermutated sequences (Umar et al., 1991; Azuma et al., 1993). However, those systems do not yet look promising.

Analysis of somatic hypermutation in Peyer's patch B cells

Recently, germinal centre B cells from Peyer's patches of non immunized animals have been shown to provide a fast and reliable source to analyse statistically significant number of clones, bypassing hybridoma production (González-Fernández and Milstein, 1993; Betz et al., 1995; Yélamos et al., 1995). PP germinal centres are found in the absence of specific immunization, probably due to stimulation by naturally occurring antigens present in the gut environment. Therefore no antigen specific immune response can be followed by this analysis. All the inconveniences described previously for the PCR approaches are also applicable to this methodology. However, when short DNA (such as V regions) fragments are amplified using high fidelity polymerases misincorporation artefacts rates are very low.

1.9 Aims of this study.

Much of the research in the last few years has focused on the identification of the DNA elements necessary to target hypermutation. Hypermutation was not affected when the V κ promoter was substituted with the promoter of the β -globin gene (Betz et al., 1994), when the V region was replaced by non-immunoglobulin sequences (Yélamos et al., 1995) or when the distance between the C κ and the 3' enhancer was shortened (Betz et al., 1994; Yélamos et al., 1995). On the other hand, removal of either the κ 3' enhancer or the κ intron enhancer/MAR causes a drastic reduction in hypermutation, which is particularly marked in the latter case (Betz et al., 1994).

There are striking parallels between the regulation of somatic hypermutation and transcription. However it is not clear whether the same regions regulate both processes. A more detailed definition of the role of the specific elements within the J-C intron and the 3' enhancer (intron enhancer, matrix attachment region, core of the 3' enhancer or its flanking regions) is needed to understand the molecular basis of somatic mutation.

The aim of this study was to establish a tighter definition of DNA regions within the κ locus important for hypermutation and to assess the role of the primary sequence and the DNA secondary structure in this process.

2.1 Materials

2.1.1 Chemicals and enzymes

Chemicals and reagents were obtained from BDH, Poole, UK or Sigma Chemical Company, Poole, UK except for those described below (analytical or 'molecular biology' grade). Radioactive nucleotides, DNA ligation system (RPN 1507), Hybond-N⁺ transfer membrane and 8.2cm Hybond-N nylon gridded membrane were obtained from Amersham International plc, Amersham, UK. Dithiothreitol was from Calbiochem, CA, USA. Sephadex G-50, Ultrapure dNTP set, Ficoll-Paque, T7 RNA polymerase and Oligolabelling kit were from Pharmacia, Sweden. 40% acrylamide/bisacrylamide stock solution and Tris buffered phenol were from Severn Biotech Ltd., UK. Hydrogen peroxide (100 volumes) and acetic acid (glacial) was from Fisons, Loughborough, UK. Ultrapure agarose, shark-tooth combs, NACS prepac cartridge and urea were from Gibco (B.R.L.), Gaithersburg, MD, USA. Saran wrap was from Dow chemical Co. USA. Acrodiscs (0.2 and 0.45µ) were from Gelman sciences, UK. Fetal calf serum was from PAA, Linz, Austria. Magic mini-preps, RNasin, RNagents Total RNA Isolation System were from Promega Co, WI, USA. Plasmid maxi and midi kit were from Qiagen Inc., Chatsworth CA., USA. Methanol and chloroform were from Prolabo, France. Ethanol was from Hayman Ltd, Witham, UK. Centricon microconcentrators (10.000 mw cut off) and Centricon filters were from Amicon Inc., MA, USA. Sequenase v2.0 sequencing kit and spermidine trihydrochloride were from United States Biochemical, Cleveland, Ohio, USA. Dialysis tubing was from Visking, UK. IPTG was from Nova Biochem, UK. Gene pulser cuvettes (0.2/0.4cm electrode gap) were from Bio-Rad Labs, CA, USA.

All restriction enzymes, DNA ligase and molecular weight markers were from New England Biolabs, Boston, MA, USA, apart from the following:

Calf intestine phosphatase, Proteinase K and Klenow fragment of DNA polymerase I were from Boehringer Mannheim, Mannheim, Germany. Taq DNA polymerase was from Cetus, Perkin Elmer, CA, USA and Promega Biotec, Madison, Wis., USA. Native Pfu DNA polymerase was from Stratagene, CA, USA. Other reagents are detailed in the appropriate sections.

2.1.2 Standard buffers

All buffers were prepared with Millipore purified water (MPW) and sterilised by filtration or autoclaving. Standard buffers and solutions are listed below.

ABTS: 3 tablets of ABTS were added to 23ml 0.1M sodium citrate and 27ml of 0.1M citric acid.

25% AMPS (w/v): 2.5g ammonium persulphate in 10ml H₂O.

rATP (10x): 10mM rATP in H₂O adjusted to pH 7.0 with NaOH.

Church buffer for hybridisation (1x): 0.5M NaH₂PO₄/Na₂HPO₄ buffer pH 7.2, 7% SDS, 2mM EDTA.

Church wash solution: 0.5M phosphate buffer, 1% SDS.

Denhardt's solution (100x): 2% (w/v) BSA, 2% (w/v) Ficoll, 2% (w/v) polyvinylpyrrolidone (PVP).

DNA gel loading buffer (6x): 0.25% bromophenol blue, 0.25% xylene cyanol FF, 15% Ficoll 400 in water.

DTT (1M stock): dissolve 3.09g DTT in 20ml of 0.01M sodium acetate, pH 5.2. Filter sterilise, aliquot and store at -20°C.

Hybridisation solution: 5x SSC, 10x Denhardt's, 1% SDS, 100mg/ml heparin.

PBS: 125mM NaCl, 8.4mM NaH₂PO₄, 16.6mM Na₂HPO₄, pH 7.2.

PEG precipitation solution: 20% PEG 8000, 2.5 M NaCl.

Sodium acetate 3M, pH 4.8.

Spermidine 100 mM, pH 7.0.

SSC (20x): 3M NaCl, 300mM sodium citrate, pH 7.0

STE: 0.1M NaCl, 10 mM Tris.HCl pH 8.0, 1 mM EDTA (pH8.0)

TAE (50x): 2M Tris-HCl, 4M sodium acetate, 20mM EDTA, pH 7.8

TBE (10x): 0.89M Tris-HCL, 0.86M boric acid, 25mM EDTA, pH 8.3

TE : 10mM Tris pH 7.4, 0.1mM (1mM) EDTA

TfBI : 30mM potassium acetate, 50mM MnCl₂ 100mM KCl, 10mM CaCl₂ and 15% (v/v) glycerol. in H₂O.

TfBII : 10mM 3-(N-morpholino)propane-sulphonic acid (MOPS) pH 7.0, 75mM CaCl₂, 10mM KCl 15% (v/v) glycerol.

Maxam-Gilbert buffer: 0.1% SDS, 1mM EDTA, 10mM MgCl₂, 500mM ammonium acetate.

Formamide dyes: 0.1% bromophenol blue, 0.1% xylene cyanol FF, 10μM EDTA in deionised formamide.

TNE: 10mM Tris pH 7.5, 5mM EDTA, 300mM NaCl.

2.1.3 Media

Media and plates for the growth of *Escherichia coli* were as follows:

Agar plates:

H-top agar: 0.8% bactoagar, 1% bactotryptone, 0.8% NaCl

Minimal glucose agar: 60mM K₂HPO₄, 40mM NaH₂PO₄, 15mM (NH₄)₂SO₄, 14μM ZnSO₄, 1μM MnSO₄, 1.2μM H₃BO₃, 0.7μM CaSO₄, 2.5μM CaCl₂, 1.8μM FeCl₃, 1mM MgSO₄, 0.002% thymine, 30mM glucose, 2% agar.

TYE agar: 1.5% agar, 0.8% NaCl, 1% bactotryptone, 0.5% yeast extract.

Media:

Tryptone broth: 10g tryptone, 5g NaCl in 1litre H₂O, pH 7.4.

2xTY liquid medium: 1.6% tryptone, 1% yeast extract, 0.5% NaCl pH 7.4.

SOB: 2.0% tryptone, 0.5% yeast extract, 0.05% NaCl, pH 7.4.

SOC: To 100ml of SOB add 5ml 20% glucose, 1ml 1M MgCl₂ and 1ml 1M MgSO₄.

2.1.4 Bacterial strains

All cloning procedures were performed with the TG1 strain (Gibson, 1984) which has the following genotype: F' traD36 lacI^q Δ(lacZ)M15 proA⁺B⁺/supE Δ(hsdM-mcrB)5 (r_k⁻m_k⁻McrB⁻) thi Δ(lac-proAB).

2.1.5 Origin of plasmids

The following plasmids were used routinely:

Bacteriophage M13mp18 and M13mp19 (Yanisch et al., 1985), pSV2-neo (Southern & Berg, 1982), pUC19 and pUC18 (Yanisch et al., 1985), pBGS19 (Spratt et al., 1986), pBluescript KS (+) (Stratagene), Lκ (Sharpe et al., 1991), LκΔB (Yélamos et al., 1995), B10513 and K472 were gifts from A. Betz, pT7Cκ was a gift from Gareth T. Williams.

2.1.6 Cell lines

All the fusions were performed using the plasmacytoma cell line NS0 (Clark and Milstein, 1981). All hybridoma cell lines were obtained from the MRC Laboratory of Molecular Biology, Cambridge. For transfections the plasmacytoma cell lines NS0 and S107 (Salk Cell Bank) were used.

2.2 Methods

Unless otherwise specified, all methods were described by Sambrook et al. (Sambrook, et al., 1990).

All centrifugations were done in either an MSE Micro Centaur microfuge (for volumes up to 2ml), IEC CENTRA MP4 (for volumes up to 50ml and speeds not exceeding 4000rpm), a Sorvall RC-5B refrigerated super-speed centrifuge (for volumes greater

than 50ml or refrigerated runs). All methods which involve the handling of bacterial cultures were performed under semi-sterile conditions following the appropriate safety codes. Cultures were incubated at 37°C with constant shaking in either a G24 environmental incubator shaker (for volumes of up to 5ml) or a controlled environment incubator shaker (for volumes up to 500ml), New Brunswick scientific, Edison, USA.

2.2.1 Construction and preparation of plasmids

Preparation of plasmid DNA

For the preparation of plasmid DNA for analytical purposes, a single bacterial colony harbouring the required plasmid was inoculated into 5 ml of 2xTY medium. This was left o.n. at 37°C with constant shaking. The culture was transferred to a 1.5ml Eppendorf tube and spun at 2000rpm for 10' and then the plasmid was prepared using magic mini-prep kit following the manufacturer's specifications. For large scale preparations the Q iagen midi (50ml) and maxi (250ml) kit was used.

Phenol extraction and ethanol precipitation

DNA samples were phenol extracted with an equal volume of Tris-buffered phenol pH 7.4. The DNA was precipitated by adding 2.5 volumes of ethanol and 1/10 volumes of 3M sodium acetate pH 5.2 and placed for 20' on a dry ice/isopropanol bath or at -20°C o.n. After centrifugation samples were washed with 75% ethanol to remove salt and then air dried and resuspended in TE.

DNA quantitation

10µl of DNA stock was diluted 1:50 in H₂O and its O.D.₂₆₀ determined using a quartz cuvette. The spectrophotometer was first zeroed on a sample containing H₂O only. O.D.₂₆₀=1 was taken to be equivalent to 50µg/ml for dsDNA, 40µg/ml for ssDNA and 20µg/ml for synthetic oligonucleotides. DNA samples with O.D.₂₆₀:O.D.₂₈₀ ratios below 1.8 contamination were repurified through phenol extraction/ethanol precipitation. DNA fragments used for the transfection of mammalian cells or for microinjections were quantified by serial dilution of samples run on an agarose minigel alongside molecular weight markers of known concentration visualised through EtBr staining.

Restriction enzyme digest and gel electrophoresis

Restriction digests were carried out in 50 to 200µl total reaction volumes, depending on the amount of DNA being digested (3 to 30µg) in the recommended buffer. A four to five fold excess of enzyme units was routinely used. The amount of glycerol in the reaction due to the enzyme added was never allowed to exceed 10% of the total reaction volume. The reactions were incubated for at least 1h. Separation of the digestion product

was achieved by horizontal gel electrophoresis using 0.7-2.0% agarose (depending of the DNA fragment size) dissolved in 1xTBE. Analytical agarose gels (20ml) were run in 1xTBE containing 1 μ g/ml ethidium bromide at 50mA, while preparative gels were made up using 100ml of agarose. These were run at 100mA.

Restriction of PCR products with two enzymes were performed using recommended buffer for the double digestion. These digests were incubated at 37°C for at least 4 hours.

Elution of DNA fragments and ligation reaction

DNA fragments were visualised after ethidium bromide staining by exposing the gel to UV light. Purification of DNA fragments was carried out by cutting the agarose gel on the right side of the band of interest. A sandwich of glass fibre filter paper and dialysis tubing was inserted into this cut. The gel was placed back into the tank after turning it 90°. Electrophoresis was then continued until the DNA had reached the glass filter. The sandwich containing the DNA was then placed into a 0.5ml Eppendorf tube with a hole in the bottom placed inside a 1.5ml Eppendorf. The liquid containing the DNA fragment was spun down into the 1.5ml Eppendorf by a quick centrifugation and the recovered material was phenol extracted and subsequently precipitated.

All ligation reactions were carried out to manufacturer's specifications with T4 ligase or a DNA ligation system (Amersham Kit). They were incubated for 5 hours or o.n. at 15°C. Religation of the vectors was avoided through dephosphorylation using calf intestinal alkaline phosphatase following the recommended reaction conditions. The ligated DNA fragments were used to transform competent bacterial cells (see below).

Preparation and transformation of competent cells

This protocol is a variant of one developed by M. Scott, Dep. of Neurology (UCSF). The transformation efficiency of these cells was 10⁵-10⁶ plaques per μ g of M13mp18. The following buffers were freshly made from sterile stock solutions before each batch of frozen competent cells was prepared:

25ml of 2xTY in a 250ml flask were inoculated with a single colony of bacteria from a minimum glucose agar plate. The culture was grown at 37°C in a shaking incubator to midlog phase (O.D.₆₀₀ = 0.2-0.8), diluted into 100ml 2xTY in a 2l flask and grown again until the O.D.₆₀₀ = 0.5-0.9. The culture was diluted once more to 500ml with 2xTY and the cells were harvested when an O.D.₆₀₀ of 0.6 was reached. The culture was then pelleted by centrifugation for 15' (4200 rpm) at 4°C. The cells were resuspended gently in 100ml of ice cold TfbI and spun again 8' (4200 rpm) at 4°C, gently resuspended in 20 ml of ice cold TfbII and aliquoted in prechilled 1.5ml Eppendorf tubes on dry ice. The aliquots were frozen in liquid nitrogen and stored at -70°C for up to 4 months.

The transformation was performed following the standard procedures. After thawing the frozen competent cells, 180 μ l of bacteria were incubated on ice with no more than 20 μ l

of DNA solution for 30'. Then they were heat shocked at 42°C for 2' and returned onto the ice. In case of M13 transformations cells were then mixed with 3ml of melted H-top agar, which had been cooled to 55°C, and poured on H-agar plates. The plates were allowed to settle and incubated at 37°C o.n. Blue/white selection was used to help identify vector-containing inserts (which will produce white colonies by disrupting the β galactosidase gene in the vector). 25 μ l of 2.5% IPTG and 25 μ l of 2% X-gal were added to the 3ml of top agar just before plating.

In the case of plasmid vectors, cells were recovered after transformation for 30' to 1 hour in 1ml 2xTY at 37°C before plating on agar containing the selective antibiotic. For pUC based vectors ampicillin (100 μ g/ml) was used. For pBGS derived vectors kanamycin (50 μ g/ml) was used. Plates were incubated at 37°C o.n.

Preparation and electroporation of 'electro-competent cells'

Electro-competent cells were prepared from a 1litre of bacterial culture in 2xTY inoculated with 1/100 volume of an o.n. culture. This was grown to about O.D.₆₀₀ = 0.5-5.7 and then chilled for 15' on ice. The culture was centrifuged at 5000 rpm for 20' and resuspended in the original volume of ice cold 1mM Hepes, pH 7.0. Cells were respun and resuspended in half of the original volume of the same buffer. The cells were then washed with 20ml 10% glycerol containing 1mM Hepes, pH 7.0, and resuspended in 2-3ml 10% glycerol (the cell concentration should be 3×10^{10} /ml). The cell suspension was frozen in aliquots on dry ice and transferred to -70°C

The electroporation was used when high transformation efficiency (10^8 - 10^9 colonies / μ g of plasmid) was required. For electroporation the DNA ligation mixture (500ng minimum) was purified through extraction with phenol and ethanol/salt precipitation and then resuspended in up to 20 μ l of H₂O.

Frozen electro-competent cells were thawed on ice. 50 μ l of cells were mixed with the DNA and transferred to a prechilled 0.2cm cuvette. The cuvette was placed on ice for 2' and finally dried before electroporation. A Bio-Rad Gene Pulser (Bio-Rad Lab. Ltd., Richmond, CA) was used to deliver 1 pulse of 2.5kV (to give 25 μ F with the pulse controller set to 200 ohms). Immediately afterwards 1ml of fresh SOC was added to the cells for resuspension, which were then incubated for 1 hour at 37°C in an incubation shaker before plating onto the appropriate plates.

Preparation of M13 templates for sequencing

Single stranded DNA from recombinants in M13mp18 (Sanger et al., 1980) was prepared from white plaques, which had previously been identified as positive for the relevant gene by colony hybridisation. These were toothpicked into 1.5 ml of a 1/100 dilution of a TG1 o.n. culture in 10ml tubes and incubated for 4.5 hours at 37°C in a shaker. Cultures were then spun at maximum speed for 5' and 1ml of the resulting

supernatant was carefully removed and mixed with 200µl of 20% (w/v) PEG 8000/2.5M NaCl. The mixture was left at RT for 10' and then spun for 10' at maximum speed. The supernatant was aspirated and the residual PEG solution was removed after a second spin. The pellet was fully resuspended in 200µl TE and then extracted with an equal volume of Tris-buffered phenol followed by ethanol/salt precipitation. The DNA was finally washed with 70% ethanol, dried and redissolved in 50µl of TE.

2.2.2 Preparation of genomic DNA

This protocol was used to prepare genomic DNA from eukaryotic cells ($1-2 \times 10^7$) and from mouse tails (approximately 1cm long) for Southern blotting or PCR.

After two washes in PBS the cells were resuspended in 1ml of Solution A (75mM NaCl, 25mM EDTA pH 8.0), then 1ml of Solution B (10mM Tris/ HCl pH 8.0, 10mM EDTA pH 8.0, freshly added 400 µg/ml proteinase K) was added. The mixture was incubated for 6h or o.n. at 55°C. The sample was extracted with an equal volume of phenol, mixing very gently to avoid shearing of the DNA, and later with chloroform. Then, 200µl of 3M sodium acetate and 5ml of ethanol were added, inverting the tube gently several times until the DNA formed a precipitate. The DNA was fished using a plastic loop and left to air dry. It was dissolved in an appropriate volume of TE. If required, the sample was incubated in 20µg/ml RNase for 15' at 37°C, to remove unwanted RNA. Repeated freezing and thawing was avoided. DNA was stored at 4°C for frequent use and at -20° C for long term storage.

Tail samples were collected from mice one week after weaning and placed in 2ml of PBS. Samples were homogenised at high speed for 30" (3x) using a Polytron Kinematica (PCU). Cell suspensions were centrifuged for 10' at 3000rpm and the pellet was resuspended in 1ml of Solution A and processed as before. When mouse liver was the source of DNA it was obtained following the same protocol, but keeping the sample in ice all time before the proteinase K digestion.

Isolation of DNA from small numbers of cells

This protocol was used to obtain DNA directly from cells sorted in the FACStar plus, for PCR. Cells (1×10^4 - 10^5) were collected into a 0.5ml Eppendorf tube and spun down in a microcentrifuge at the low speed setting for 2'. The supernatant was carefully removed and the cells were resuspended in 10-20µl of water (depending on the number of cells) by vortexing, frozen immediately in liquid nitrogen and kept at -70°C. The DNA was isolated by treating the thawed cells with proteinase K (100µg/ml) in 1x PCR buffer/0.5% Tween 20 (20-40µl final volume) and incubated for 1h at 56°C. The sample was heated to 95°C for 30' to inactivate the proteinase K and kept at -20°C.

2.2.3 Production of P³² radiolabelled probes

Radioactive labelling of oligonucleotides

Oligonucleotides were synthesised by Terry Smith (MRC, LMB, Cambridge, UK) using an Applied Biosystems AB380B DNA synthesiser. The oligonucleotides (which do not possess 5' phosphate group) were labelled by setting up a standard phosphorylation reaction with 10.2µl TE, 4.3µl H₂O, 2.5µl (10x) kinase buffer, 5µl of (γ-³²P)-dATP (10µCi/µl), 1µl of T4 polynucleotide kinase and 3µl of oligonucleotide at 10pmoles/µl. The reaction was carried out for 1h at 37°C and stopped with 100µl of TE. The labelled probe was separated from the radioisotope non incorporated by chromatography over a Sephadex G50 column before it was used for hybridisation.

Radiactive labelling of DNA fragments

Gel purified DNA fragments were labelled with (α-³²P)-dCTP (10µCi/µl) by primer extension using a Pharmacia oligolabelling kit, according to the manufacturer's specifications. The probe was purified by Sephadex G50 chromatography before it was used for hybridisation.

Spin column preparation and use

Glass wool was siliconized by putting the glass wool in a beaker in a hood, adding silane solution, covering the beaker with aluminium foil and leaving for 4h to dry in a oven at 150°C. A 1ml plastic syringe was plugged with siliconized glass wool and a 0.9ml Sephadex G50/TE (pH8) spin column prepared in the following way. Sephadex G50 was poured (carefully to avoid bubbles) into the column to the 1.3ml mark. The column was then placed into a 15 ml conical bottomed tube with an Eppendorf inside and spun for 2' at 1500rpm. The column was repeatedly topped up and spun until the final packed volume was 1ml. The sample was then applied to the top of the column. The eluate was collected in a fresh Eppendorf at the bottom of the 15ml tube.

2.2.4 Screening of bacterial clones

The Buluwela protocol for screening bacterial colonies or M13 plaques (Buluwela, et al., 1989) was routinely used. 0.45µ, 82mm circular Hybond-N filters were used to lift bacterial colonies or M13 plaques and markings made with a needle. After the standard fixation procedure filters were prehybridized for 10' at 65°C and then hybridised for at least 4h at 65°C. First washes were carried out in 4xSSC at RT, further washes in 2xSSC/1%SDS for 15' increasing the temperature up to 5°C below the melting temperature for the specific oligonucleotide, until the negative control could no longer

be detected. The filters were then wrapped in Saran and exposed 1-2h or o.n. placing Stratagene glowing markers to allow orientation.

2.2.5 Southern Blots

For each digest 20µg of DNA were used and digestions were carried out o.n. Then the samples were phenol extracted and ethanol/salt precipitated. After resuspension the digestion products were separated on a 250ml 0.7% agarose (dissolved in 1xTAE) gel which was run o.n at 45mA, side by side with the appropriate mw markers. The gel was run until the fast blue was out of the gel and stained for 15' with EtBr. A photograph was taken with a fluorescent ruler by the side of the gel to mark the positions of the markers, also, an exact replica of the gel was drawn on a transparent sheet for an easier location of the relevant bands. The presence of satellite DNA indicated that the digestion had gone to completion. The gel was placed in 0.25N HCl until the dyes changed colour and then for another 10'. Then the HCl was washed away with distilled H₂O and the DNA was denatured for 15' in 0.5M NaOH. Hybond N⁺ membrane was cut to size, soaked in 0.5M NaOH for 5'. Two 0.5M NaOH soaked strips of 3MM paper were placed on a glass plate so that they overhung into a tray full of 0.5M NaOH. The gel was placed on top of the glass plate and the Hybond N⁺ membrane was then placed on top of the gel and rolled to remove air bubbles. A piece of 0.5 M NaOH soaked 3MM paper was then placed on top of the Hybond N⁺ membrane and the paper strips on the glass plate still exposed to the air were covered with parafilm to prevent evaporation through these areas. Three more dry 3MM pieces were placed on top of the gel letting them get wet and taking care to avoid air bubbles. A large stack of paper towels was then positioned on top of the 3MM paper. The gel was left to blot for 16-24 h at RT with a change of paper towels and several top ups 0.5M NaOH as required. Before removing the membrane from the gel, the position of the DNA slots was marked. The filter was then peeled from the gel and neutralised in 1.5M NaCl/0.5M Tris-HCl pH 7.4 for 15' at RT, then the DNA was cross linked to the membrane in a Stratalinker (Stratagene) at the standard settings before storage at RT in foil. The membrane was then placed in a glass rolling bottle and prehybridised for 10' at 65°C in 10ml Church hybridisation buffer. The probe was boiled for 5' and put on ice for 10'. This was then added to new preheated hybridisation fluid in a Falcon tube. The prehybridisation fluid was then discarded and the probe/hybridisation fluid mixture added to the membrane. After o.n. hybridisation at 65°C, the radioactive hybridisation fluid was decanted into a 50 ml Falcon tube for future use (such mixtures were used up to 5 days after initial preparation, boiled before re-use). Filters were then washed 2x for 15' at RT with Church wash buffer, 2x for 20' at 50°C monitoring the counts until the background was reduced to a minimum. The filter was then wrapped in Saran and put into a cassette with two Fuji RX medical X-ray film (Fuji Photo Film Co. Ltd., Japan) (one on each side of the filter) for autoradiography o.n.

at -70°C using an intensifying screen. The developing process was carried out automatically using a Kodak RP X-OMAT Processor.

2.2.6 DNA sequencing

All sequencing of ds and ssDNA was done using the Sanger dideoxy chain termination method (Sanger et al., 1977). A Sequenase Version 2.0 sequencing kit (United States Biochemical) was used for all sequencing reactions following the manufacturers instructions, with the exception that extension was done at 18°C. and termination was at 42°C. Sequencing was carried out in microtiter round-bottomed plastic plates. Only 1.5µl from the sequencing mix were loaded to improve the resolution of the bands in the gel.

Preparation of acrylamide gradient sequencing gels

To produce a 6% acrylamide gradient sequencing gel the following stock working solutions were made:

5xTBE mix: 15ml 40% acrylamide, 50ml 10xTBE, 46g urea and a pinch of bromophenol blue.

0.5xTBE mix: 75ml 40% acrylamide, 25ml 10xTBE, 230g urea and H₂O to top up to 500ml.

The two solutions were filtered through a 0.45µm Nalgene filter and kept at 4°C for up to 1 month. Two glass electrophoresis plates for sequencing (one with a bevelled end) were washed with detergent, rinsed with H₂O and then with ethanol. The one with a bevelled end was siliconated. Spacers were put at either side between the plates and the plates were then taped together tightly with water proof tape. 13µl of 25% AMPS was added to 7.5ml of the 5xTBE and 88µl of 25% AMPS to 50ml of the 0.5xTBE. Finally 13µl of TEMED was added to the 5xTBE and 88µl of TEMED to the 0.5xTBE. 6ml of 0.5xTBE was taken up in a 25ml pipette and 7ml of 5xTBE (blue) was taken up in the same pipette. The gradient was achieved by allowing three air bubbles to pass through the pipette. The electrophoresis plates were held at 45° and the gradient poured down the middle of the plates followed by the rest of the 0.5xTBE solution. The shark-tooth comb was placed in an inverted position to create a mould and the plates clipped together. The gel was left to set for a minimum of 1h. Before use, the comb was extracted, the surface of the gel edge was cleaned with H₂O and the shark-tooth comb inserted to touch the surface of the edge of the gel. Gels were pre-run for 30' before the samples were loaded. Gels were run at 37 W for approximately 3h in 1xTBE buffer and fixed for 15' in 10% methanol/10% acetic acid, blotted on to 3MM paper and dried under vacuum for 45' at 80°C. Autoradiography was performed o.n. at RT.

2.2.7 Polymerase chain reaction (PCR)

PCR conditions were optimised for each set of primers. Primers for PCR were adjusted to 10pmoles/ml. It was sometimes necessary to gel purify some primers (see below). Reactions were performed in 0.5ml Eppendorfs and overlaid with light mineral oil followed by heating to 94°C for 10' in a thermal cycler (PHC-2 from Techne or Biometra PCR machine) to denature the DNA. The enzyme was added under the mineral oil and the mixture was then thermal cycled (30x). Unless otherwise noted, the cycle was: denature for 2' at 92°C, anneal for 1.5', and extended for 2' at 72°C. An initial annealing temperature was selected based on the melting temperature of the oligonucleotide primers and it was decreased 5°C if the amplification failed. An extra extension step was carried out for 10' at 72°C.

Either 5pg of plasmid DNA or 0.1-1µg of genomic DNA was employed for each amplification. When cell lysates (from Peyer's patches cell sorted cells) were used 3, 6 and 10µl from the mixture were utilised to standardise the PCR conditions. From 1992 to 1994 the Cetus Taq polymerase was used, the standard reaction was performed in a total volume of 50µl following the manufacturers conditions. From 1994 Native Pfu DNA polymerase (Stratagene) was used since the frequency of errors introduced is considerably lower than the other polymerases. In this case the amplification was performed in a final volume of 100µl according to the enzyme manufacturers specifications. Usually multiple amplifications were performed at the same time for a single sample.

Contamination during PCR from previously cloned DNA was a major concern, particularly when PCR from PP cells from transgenic mice was carried out to analyse somatic hypermutation. To avoid contamination, the following measures were taken: 1) gloves were worn at all the time, 2) PCR reactions were set-up in a bubble chamber using a set of separate pipettes kept exclusively for PCR, 3) all reagents were aliquoted, used once and discarded, 4) a negative control replacing the DNA with H₂O was always included to ensure that inadvertent contamination had not occurred, 5) plugged pipette tips were always used, 6) gel boxes and combs were depurinated o.n. with 0.25M HCl

PCR screening of recombinant clones

PCR was used to analyse recombinant clones for the presence of the correct insert, after a DNA ligation, directly from bacterial colonies (Gussow and Clackson, 1989). A PCR reaction mixture was made up containing Promega Taq polymerase buffer, 200µM dNTPs, 20pmol of each of the oligonucleotide primers flanking the insert of interest and 0.3µl of Taq polymerase per 20µl of reaction mixture. Template for amplification was transferred into the tube by touching the bacterial colony with a small plastic inoculating loop. The reaction mixtures were overlaid with 50µl of light mineral oil and denatured at 94°C for 10' followed by thermal cycling using the specific conditions for the set or

primers chosen. 20 cycles were performed usually. Clones containing the correct insert were identified either by the size of the insert, or when necessary, by digesting the amplified DNA with restriction enzymes and examining the restriction fragment sizes on an agarose gel.

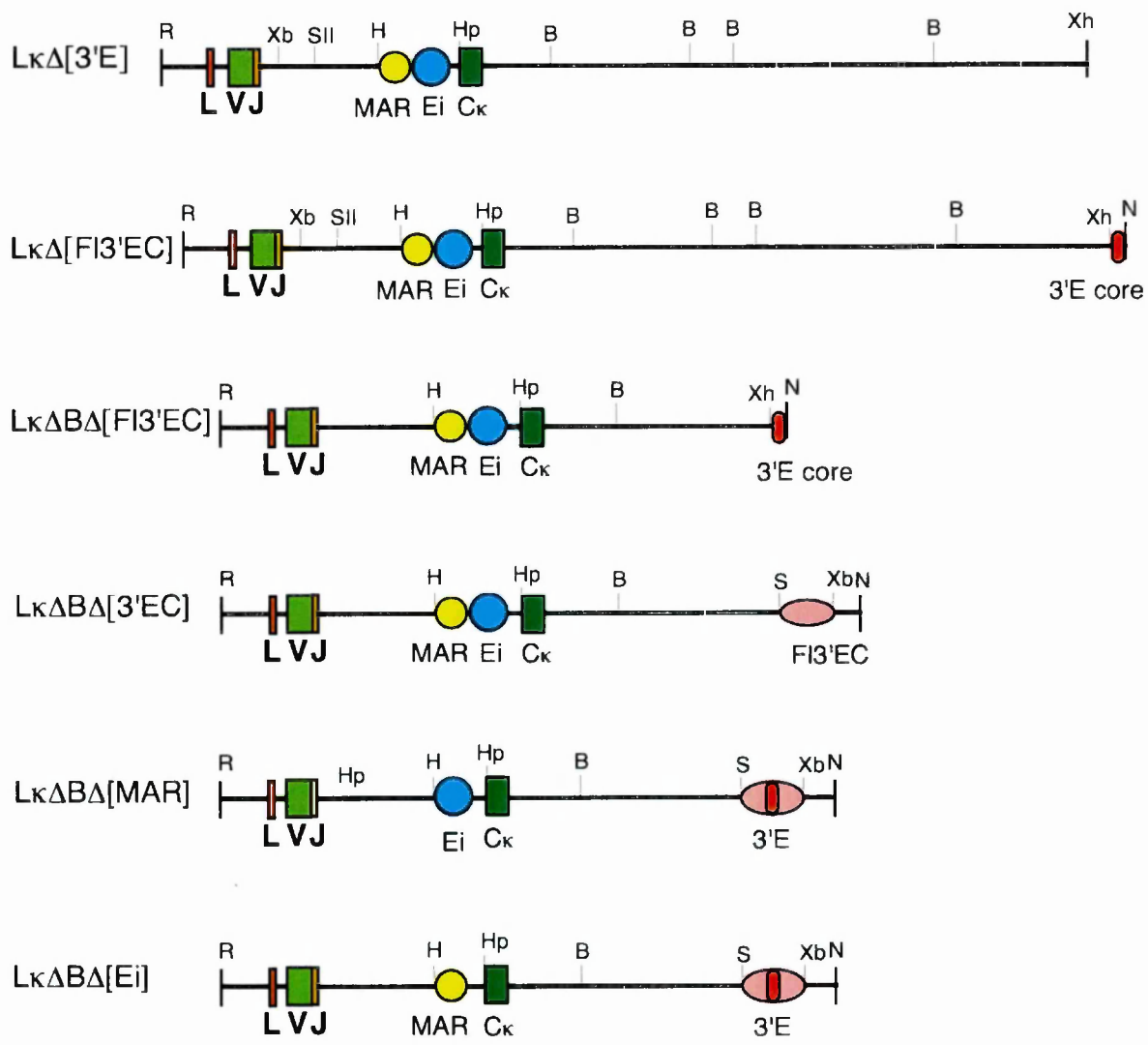
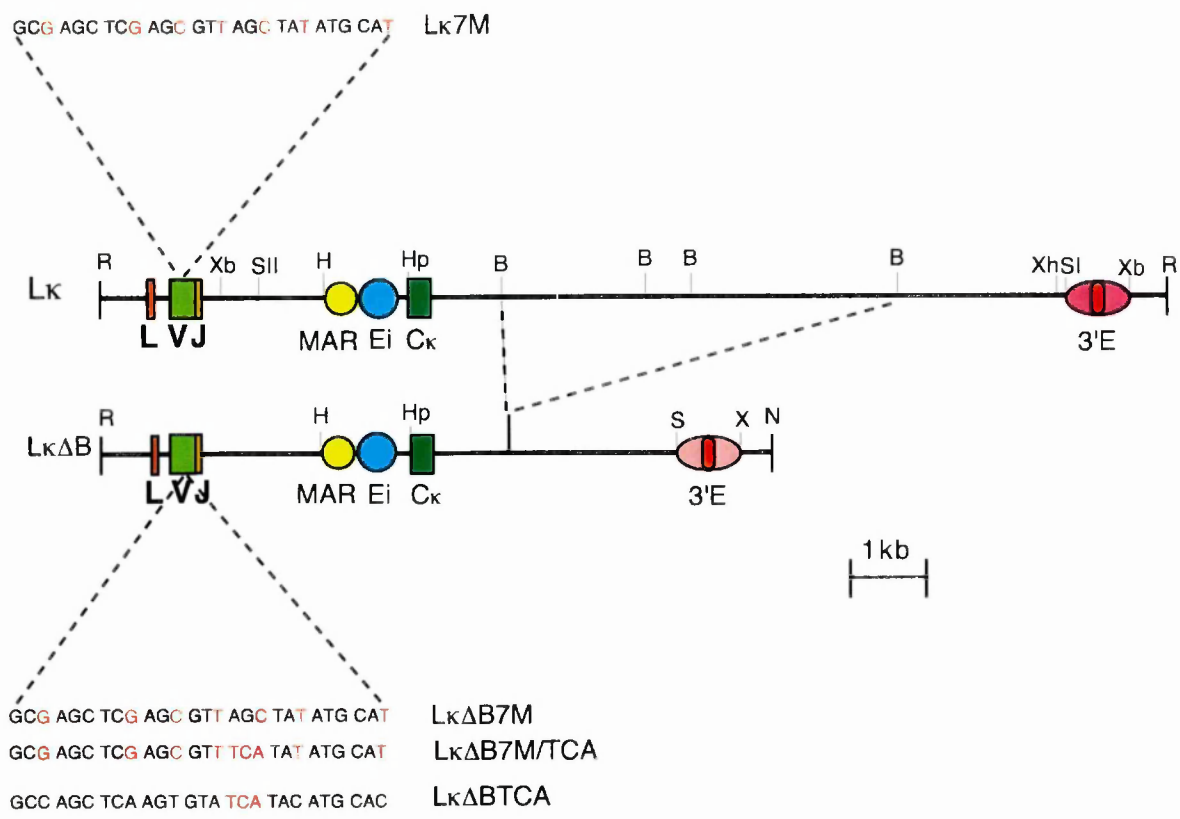
Purification of synthetic oligonucleotides

Small glass plates (40x20cm) were cleaned and prepared as for sequencing gels. The percentage of acrylamide used depended on the size of the oligo being purified. A useful guidance follows: for 25 nt or less 20% acrylamide gels were used; for 25-40 nt 16% acrylamide gels were used. A 45ml 7M urea 1xTBE acrylamide gel mixture was prepared and 90µl of 25% AMPS and 90µl TEMED added to this immediately before pouring. Gels were allowed to set for 1h 8µl of neat oligo (approx. 4 O.D.₂₆₀ units) were mixed with 8µl of loading buffer and heated at 80°C for 5' to disrupt any secondary structure before being loaded. The gel was run in 1xTBE buffer at 34W, until the fast blue had run approximately two thirds of the length of the gel. The position of the oligo was estimated from the position of a control oligo of known length. The gel was then transferred onto Saran and the bands UV shadowed using long wave UV (the DNA absorbs UV and so appears black against the uniform fluorescent background of a F254 silica gel 60 fluorescence indicator chromatographic plate). The highest mw band representing the undegraded oligo was circled with a felt tip pen and cut out. The gel slice was then placed in a Eppendorf tube and reduced to small pieces with a scalpel, then 200µl of H₂O was added and this was shaken 5h or o.n. at 37°C. After a short centrifugation the supernatant containing the eluted oligonucleotide was ethanol/salt precipitated and resuspended in 50µl of H₂O, its O.D.₂₆₀ determined and then adjusted to 10pmoles/µl. Great care was taken throughout the whole manipulation to avoid contamination of the oligonucleotide with cloned DNA.

2.2.8 Construction of plasmids

Lκ as basis for further constructs

All the constructs used in this work are based on the original Lκ construct assembled by M. Sharpe (Sharpe et al., 1990), encoding a rearranged mouse (BALB/c) Vκ gene, specific for the hapten phenyl-oxazolone (VκOx1), linked to a rat Cκ gene. The presence of the rat constant region meant that the transgene could be distinguished from the endogenous κ at the protein level using anti-rat κ antibodies and at the nucleic acid level by a ribonuclease protection assay. The transgenes consist of mouse κ sequences except



for the region from the Hpa I site in Cκ to the Bgl II site immediately 3' of Cκ, which has been substituted by sequences from the LOU rat strain and a deletion of 27nt flanking the Xba I site downstream of Jκ5. This difference allows unambiguous discrimination of the transgene from the endogenous loci by differential hybridisation with the oligonucleotides P1400 and AFR3 (see Appendix II), which hybridise with the wild type and deleted forms respectively. The Lκ transgene is known to be a good substrate for the hypermutation process (Sharpe et al., 1991) as well as a short version of this construct LκΔ[3'F] (Betz et al., 1994). For cloning conveniences LκΔB, a short version of the Lκ where the DNA fragment between the BamH I sites 3' of the Cκ was deleted (Yélamos et al., 1995), was also used to make other constructs.

The construction of individual plasmids used to create transgenic mice and those used in the transfection studies are described in the following sections. The structure of the transgene in those constructs are shown in the Fig. 2.1.

Figure 2.1 Structure of transgenes (opposite page).

The structure of the different transgenes used are depicted. In the structure of the transgenes, carrying the CDR1 mutated, the Vκ segment containing the mutations is depicted. The nature of the mutations (which are shown in red) in each transgene are indicated. The restriction endonuclease cleavage sites are abbreviated: R, EcoRI; B, BamHI; Xb, XbaI; Xh, XhoI; SI, SacI; SII, SacII; H, HindIII; Hp, HpaI; N, NotI.

2.2.8.1 The L κ -CDR1 mutated constructs

A series of oligonucleotides was synthesised in order to replace the original DNA sequence at the CDR1 of V κ Ox gene without changing the amino acidic sequence (see Appendix I A). This was achieved by designing them in pairs: one of them is the coding region carrying the mutations and the other is its complementary sequence. When they were hybridised they acted as a double stranded DNA synthetic fragment with cohesive ends ready to ligate.

50pmol of each oligonucleotide was phosphorylated using the standard protocol with the PN Kinase, followed by heat inactivation and ethanol precipitation. Each oligo was resuspended in 10 μ l of H₂O and mixed with its complementary. The mixture was boiled for 5' in a water bath and hybridisation was allowed to occur for 40' by letting the water bath cool to room temperature. The tubes were placed on ice and 4 μ l of the mixture was used to ligate with 2 μ g of vector using the Amesham kit. After o.n. incubation at 15°C half of the ligation was used to transform and the recombinant clones were analysed by digesting with restriction endonucleases and by sequencing the V κ Ox gene with V κ Ox212 oligo (Appendix II).

Firstly the PstI/KpnI fragment of the plasmid pB10513 (pUC19 containing the EcoRI-HindIII L κ fragment) was replaced by the synthetic fragment composed of the oligonucleotides CDR1M3 and CDR1M4, which were degenerated in several positions allowing the formation of different variants of mutants. The individual clones were selected from the "library" by double stranded sequencing with V κ Ox212 oligo and by restriction enzyme digestion, since some of the mutations were meant to generate new restriction enzyme cuts in order to facilitate the identification of each copy (see in Appendix I for the list of the selected clones carrying the mutagenised CDR1). The fragment EcoRI-SacII from three mutants (pBEA25, pBEA28 and pBEA31) were used to replace the original L κ sequence in the plasmid pK471 (pUC19 containing the EcoRI/BamHI L κ fragment) since EcoRI-BamHI fragments carrying mutations were required for microinjection.

The oligonucleotides CDR1M1 and CDR1M2 (carrying seven mutations) were employed to replace the original fragment in pB10513 and pK471. The mutant EcoRI/SacII fragment cloned into pB10513 (pBEA7) was used to replace the original sequence in the L κ whilst the mutant EcoRI-BamHI fragment cloned into pK471 (pKBEA7) was used to replace the original sequence in the L κ B Δ B (derivative of L κ Δ B with the BamHI site in the pSV2neo destroyed to facilitate the cloning procedures). The L κ 7M, and the L κ Δ B7M were used to purify DNA fragments employed in the egg injection and they were used extensively as a basis for further constructs, since they carry three new restriction sites in the CDR1 region making the identification of new L κ derivatives simpler.

Another CDR1 mutated construct (LκΔBTCA) was made replacing the PstI-KpnI fragment of pB10513 with a synthetic analogue composed by the oligonucleotides CDR1M7 and CDR1M8. The plasmid pBEA60 was used then, to take the fragment EcoRI/SacII, containing the "TCA" mutation, and ligated into the LκΔB.

The last of the CDR1 mutants (LκΔB7M/TCA) was generated by replacing the XhoI/KpnI fragment from pBEA6 with the synthetic analogue made with the CDR1M5 and CDR1M6 oligo combination. This construct (pBEA72) was used to obtain the EcoRI/SacII fragment containing the "7M/TCA" modifications and ligated back into its original place in the LκΔB construct.

2.2.8.2 The κ 3' enhancer dissection constructs

Three independent deletions were produced in the κ 3' enhancer region (Meyer and Neuberger, 1989). One includes the whole 3' enhancer, the other its core (Meyer et al., 1990; Pongubala and Atchison, 1991) and the third, the flanking regions of the 3' enhancer core.

LκΔBA[3'E]

The construct carrying the complete deletion of the 3' enhancer region (LκΔBA[3'E]) was made by digesting the plasmid LκΔB with XhoI and NotI, followed by filling in the protruding 5' ends with the Klenow fragment of the DNA polymerase I. In this way the NotI site was destroyed and the XhoI was reconstituted after religation.

LκΔ[F13'EC] and LκΔBA[F13'EC]

The deletion of the flanking regions of the 3' enhancer core from position 12580 to 13058 and from 13200 to 13966 (Appendix II B) were made by PCR. The amplification of the core enhancer was made with the oligonucleotides DINO1 and DINO2 (Appendix II), which introduced a XhoI and a NotI site (respectively) in each side of the core. The 142bp PCR product was NotI/XhoI digested and then ligated back, into its original place in the Lκ and LκΔB constructs.

LκΔBA[3'EC]

The LκΔBA[3'EC] construct is a derivative of LκΔB which lacks κ 3' enhancer core. The first step of the construction was to clone the 1.3kb XhoI-NotI fragment containing the κ 3' enhancer into pBluescript II KS (-). This plasmid (pBDino5) was NcoI/BspMI digested and the protruding 5' ends filled in with the Klenow fragment of the DNA polymerase, followed by self-ligation. The insert of this plasmid (pBDino5.11) was sequenced with the M13-20 primer (see Appendix II A) to confirm that it carries the 177bp deletion (from position 13072 to 13249, see Appendix II B) which includes the 3'

enhancer core. The XhoI-NotI fragment of pBDino5.1 was then ligated back, into its original place, in the LκΔB construct.

2.2.8.3 The κ intron enhancer deletion construct (LκΔBΔ[Ei])

The LκΔBΔ[Ei] construct is a derivative of LκΔB carrying a deletion in the κ intron enhancer region. (Queen and Baltimore, 1983 ; Picard and Schaffner, 1984a). The deletion includes basically the core enhancer (Fulton and Van Ness, 1994) (see Appendix II B).

The first step in the construction was the subcloning of the Lκ EcoRI-BamHI fragment into the Km resistant vector pBGS19.5(+) (which HindIII site was destroyed). In parallel the Lκ HindIII-HpaI fragment, carrying the κ MAR/Ei elements was subcloned into pBGS19(+) HindIII/HicII digested. Thus, this plasmid (pMED3) was used to purify the 535bp AflII-HincII fragment which, together with the PCR product with the oligonucleotides MAR1 and MAR2 (containing the MAR region; see Appendix II) digested with HindIII and AflII, were assembled by a three way ligation with the vector pBGS-Lκ[E-B], HidIII/HpI digested.

The clones carrying the intron enhancer deletion (from position 1932 to 2306, see Appendix II B) were identified by PCR screening with the oligonucleotides MAR1 and MARBSEQ (see Appendix II), since the positive clones gave a 0.457kb amplification product and the negatives a 0.831kb. In addition the clones were sequenced to confirm that the assembly was correct. Finally the EcoRI-BamHI fragment from pBGS-Lκ[E-BΔ[IE]] was ligated back into its original place in the LκBΔB7M.

2.2.8.4 The MAR deletion construct (LκΔBΔ[MAR])

The MAR region (Cockerill and Garrad, 1986) was removed after amplification with the oligonucleotides IE1 and RCKN127 (see Appendix II). The 1.020kb PCR product obtained was digested with Hind III and Hpa I, and ligated in place of the equivalent Lκ fragment, in the pBGS-Lκ[E-B] vector. The clones carrying the MAR deletion (from 1572 to 1863, see Appendix II B) were identified by PCR screening with the oligonucleotides IE5 and IEBSEQ (see Appendix II A), since the positive clones gave a 0.597kb amplification product and the negatives a 0.888kb. Some positive clones were sequenced with MAR2 primer (see Appendix II) to confirm that no mistakes were introduced during the replacement. Lastly the pBGS-Δ[MAR] was digested with EcoRI and BamHI and the Lκ-Δ[MAR] segment was ligated in place of the equivalent LκBΔB-7M fragment.

2.2.9 Cell culture techniques

Freezing and thawing cell lines

To freeze cells, 5×10^6 cells or more were taken from a healthy logarithmic culture and spun at 1200rpm for 5'. The supernatant was aspirated and the pellet resuspended in 1ml of sterile freezing medium (90% FCS, 10% DMSO at 4°C) and kept on ice. The cells were transferred to 2ml freezing vials, put at -20°C for 2h and then transferred to -70°C. After a few days, the cells were transferred for long-term storage under liquid nitrogen. To thaw cells, vials were warmed quickly to 37°C in a water bath until only a little ice remained. Cells were washed in 10ml of DMEM/20% FCS and resuspended in 1.5ml of the same medium and seeded into 24 well plate by making serial dilutions. After some days cells were transferred into flasks with 10% FCS/DMEM. All cell lines and hybridomas were grown in 10% FCS/DMEM at 37°C in humidified incubators under 5% CO₂.

2.2.10 Isolation and analysis of Peyer's patches cells

Isolation of cells

Unless otherwise specified all transgenic mice used to collect Peyer's patch cells were between four and five months old. PP were recovered by dissection from the small intestine. Single cells suspensions were prepared by pressing PPs through a nylon mesh. The cell suspension was centrifuged and washed (2x) in 15ml cold DMEM. Clumps of dead cells were eliminated with a Pasteur pipette after each wash.

Fluorescence activated cell sorting

Cells were resuspended in 250µl DMEM and incubated with PNA FITC labelled (Dil. 1/60) and mAb to mouse CD45R (anti B220) PE labelled (Dil. 1/250) for 30' on ice. The labelled cells were then washed with 15ml of cold medium and resuspended in 3ml of cold PBS. The cell suspension was finally passed through a nylon mesh to avoid cell clumps which could block the sorting machine.

B220⁺ PNA^{high} and B220⁺ PNA^{low} cells were sorted on a FACStar^{plus} (Becton Dickinson) sorter by David Gilmore and Andrew Riddell.

2.2.11 Transfection of mammalian cells and analysis of gene expression

Generation of stable transfectants

Stable transformants were generated by electroporation as described by Potter et al. (1984) The DNA (50µg) was linearized with PvuI, phenol extracted, ethanol

precipitated, and resuspended in 50µl sterile PBS. Approximately 1.5×10^7 cells were spun down, washed in PBS and resuspended in 200µl of PBS. The DNA and the cells were placed in to a sterile electroporation cuvette (0.4cm gap), mixed by flicking and placed on ice for 20'. A Bio-Rad Gene Pulser (Bio-Rad Laboratories Ltd.) was used to deliver a pulse of 0.2kV (960µF, $t = 50\text{ms}$) to the cuvette, which was then placed on ice for another 20'. The cells were seeded into 24 well plates using DMEM/10% FCS. The transformants were selected 24h after transfection by resistance to G418 at an initial concentration of 0.5mg/ml and this selective concentration was increased to 1mg/ml (S107) or 2mg/ml (NS0) six days after transfection.

Pools from ten different well established clones were used for each transfectant line in order to purify the cellular RNA. The RNagents Total RNA Isolation Kit (Promega) was used following the manufacturer's instructions. Volumes were scaled down when it was necessary.

Ribonuclease protection assay

The radioactively labelled RNA probe for detecting rat Cκ was generated by T7 RNA polymerase transcription of the plasmid pT7Cκ, a Bluescript vector into which a 300bp BamHI fragment overlapping the start site of the rat Cκ has been cloned. In order to obtain the antisense probe the plasmid was linearized so that a transcription from the T7 promoter will generate an RNA probe of 360bp that is complementary to the Cκ transcript. After hybridization of the probe and the total RNA, a double stranded RNA molecule is formed with a 5' and a 3' overhang. Treatment with RNAase removes these overhangs and, generates a fragment of 300bp which can be visualised by electrophoresis and subsequent autoradiography.

Transcription from the linearized plasmid with HindIII was carried out using 2µl 10x transcription buffer (400mM Tris pH 7.5, 20mM spermidine, 60mM MgCl_2), 1µl rNTP solution (10mM each), 1µl 250mM DTT, 1µl RNasin, 5µl [$\alpha^{32}\text{P}$] UTP (10µCi/µl) and 31.5 units of T7 RNA polymerase in a total volume of 20µl for 30' at 37°C. After the incubation 10µl formamide dyes were added. The probe was purified on a 4% denaturing polyacrylamide gel as described and eluted in 1ml Maxam-Gilbert buffer 3h or o.n. at 37°C.

For the ribonuclease mapping, the radioactive probe ($5 \times 10^5\text{cpm}$) was added to 10µg of RNA which was then ethanol/salt precipitated, washed, vacuum dried and resuspended in 30µl of hybridization buffer (80% formamide, 0.4M NaCl, 10mM PIPES pH 6.4, 1mM EDTA). Samples were denature for 10' at 85°C in a water bath followed cooling the water to 46° and hybridization was allowed to occur o.n. Then 300µl of ice cold RNase mix (1µg/ml RNase-T₁, 40µl/ml RNase A in TNE) was added and single stranded RNA was digested by incubation at 30°C for 30'. To stop the reaction 20µl of

20% SDS and 7.5µl of proteinase K (20mg/ml in TNE) were added and incubated at 37°C for 15'. Samples were treated with phenol/chloroform/isoamyl alcohol, ethanol/salt precipitated, washed and dried before separated on a 4% denaturing polyacrylamide gel. A HaeIII digest of φX174 which was labelled with [$\gamma^{32}\text{P}$] ATP served as size markers. The gel was fixed, transferred, and dried as described before and an autoradiographic image was obtained after several days of exposure at -70°C.

Northern blot analysis

The cytoplasmic RNA (~10 µg) was fractionated in 1% formaldehyde/agarose gel (Lehrach et al., 1977). The gel fractionated RNA was transferred by capillary forces to Hybond N⁺ membrane in 20xSSC buffer and was then UV cross-linked to the membrane. The membrane was pre-hybridized for 30' in Church hybridization buffer and hybridized with radioactively labelled oligonucleotides in the same buffer overnight, at 42°C. The blot was washed twice for ten minutes in Church wash buffer.

The oligonucleotide used for the detection of the transgene (RATCκ18) was specific to rat sequences within the constant region on the transgene, whilst the oligonucleotide used as a control for mRNA loading (MOUSEACT) was specific to β-actin.

2.2.12 Enzyme-linked immunoabsorbent assay

Inhibition assays for detection of transgenic κ expression

This assay was developed by M. Sharpe to screen transgenic mice carrying a rearranged mouse κ light chain tagged with the constant κ region of the Lou rat. This allows easy serological detection of the transgenic light chain in the serum of the transgenic mice. This assay was also used to titrate supernatants of transfected S107 and NS0 cells and hybridomas. The strategy is based on an inhibition assay in which the test serum is incubated with an anti-rat κ reagent and the remaining free antibody is quantitated. Serum containing transgenic light chains remain uncoloured whilst serum lacking transgenic light chains resulted in a positive signal. The assay relies on the amplification step introduced with the biotin-streptavidin developing system. Falcon flat-bottomed microtiter plates were coated with 5-10 µg/ml of an irrelevant rat antibody (in this case P4C1/13-11, fraction I) and left at 4°C o.n. This was then blocked at least for 1h at 37°C with PBS/5%FCS or o.n. at 4°C and stored for up to 3 months in PBS, 5% FCS, 0.01% azide in a moist chamber. Two dilutions of test serum (1:50 and 1:250) were incubated at 4°C o.n. with an equal volume of a 1:7000 dilution (of 1mg/ml) of mouse anti-rat κ antibody MARK-1 (Serotec). Several serial dilutions were made when quantitative analysis were performed. The mixture then was transferred to the plate coated with rat Ab and incubated 90' at 37°C. The plate was then washed 3x in PBS and the amount of

MARK-1 Ab bound was assayed by incubation with biotinylated sheep anti-mouse antiserum (1:6000) (Amersham) for 45' at RT. The plate was washed 3x in PBS and of streptavidin-HRP complex (1:2000) (Amersham) was added for 15' at RT and then washed 2x in PBS, before adding the substrate solution (0.55mg/ml ABTS, 0.001% H₂O₂, 0.1M citrate buffer pH 4.3). The optimal coloration was achieved, approximately after 10', and the reaction was stopped by adding 50µl of 0.1M citric acid, 0.01% sodium azide per well. The O.D. was read in an automatic plate reader (Titertek Multiscan) at 414 nm. For quantitative ELISA serial dilutions of purified P4C1/13-11(F-I) or Y3 (Bence-Jones protein) of known concentration was used.

Inhibition assays to quantitate mouse κ expression in supernatant of stable transfected S107 cells

Falcon flat-bottomed microtiter plates were coated with 10 µg/ml of Ox14-BSA and left at 4°C o.n. This was then blocked at least for 1h at 37°C with PBS/5%FCS or o.n. at 4°C and stored for up to 3 months in PBS, 5% FCS 0.01% azide in a moist chamber. In parallel several serial dilutions were made of test supernatants were incubated at 4°C o.n. with an equal volume of a 1:800 dilution of anti-mouse supernatant MCC10 (this is a bispecific Ab that recognises mouse κ and horse radish peroxidase).

The next day, the plate was washed 3x in PBS and then incubated with a 1:1000 dilution of a mouse κ anti oxazolone NQ10.12.5.17 supernatant for 1h at 37°C. The plates were washed again, and the mixture containing the samples was transferred to the plate and incubated 90' at 37°C. The plate was then washed 3x in PBS and the amount of MCC10 Ab bound was assayed by incubation with Peroxidase (1:50) for 1h at 37°C. The plate was washed 3x in PBS and the substrate solution (0.55mg/ml ABTS, 0.001% H₂O₂, 0.1M citrate buffer pH 4.3) was added. The optimal coloration was achieved, approximately after 10', the reaction was stopped by adding 50µl of 0.1M citric acid, 0.01% sodium azide per well. The O.D. was read in an automatic plate reader (Titertek Multiscan) at 414 nm. Serial dilutions of purified 9E10 mouse Ab of known concentration was used as standard curve.

2.2.13 Transgenic mice and hybridomas

Preparation of the DNA fragments for microinjection

The constructs were digested with the appropriate restriction enzymes and the vector-free fragments were purified from the agarose gel as described before. The eluted DNA was chromatographically purified with NACS columns, following the manufacturers instructions. The eluted DNA was ethanol precipitated and washed several times until the white pellet can no longer be seen, then the DNA was dissolved in 20µl of MPW and quantitated as described before. Once the DNA concentration was known, the fragments

were either injected by themselves at the suitable dilution or a mixture of several fragments was prepared at equimolar ratio and a new quantitation, now of the DNA fragment mixture, was made. The microinjection was carried out by T. Larson and J. Jarvis.

Founder animals were identified by Southern blot analysis of tail's DNA and by serological detection of the transgenic κ light chain in the serum of mice. Founders were bred with F1 mice (CBA x C57BL/6J).

Immunisation protocols

All immunisations were performed using the same standard protocol. For primary immunisation 100 μ g of aluminium-potassium sulphate precipitated antigen (phOx-CSA) was injected intraperitoneally together with 10^9 units of heat inactivated *Bordetella pertussis* (Wellcome, UK). Secondary immunisations were done six weeks later with an intra-venous injection of 100 μ g of soluble antigen. Three days after this the splenic lymphocytes were isolated and used for fusion.

Hybridomas were established from immunised adult mice by fusion of spleen cells with the plasmacytoma cell line NS0 (Galfré & Milstein, 1981). All fusions and cloning in soft agar were made by Richard Pannell.

2.2.14 Computer analysis

Computer analysis of sequences was performed using the Staden programs on an Aliant mainframe computer. The program AFRICA used for this task is an extension of the program NIPF written by Roger Staden. Further data handling was carried out on MacIntosh personal computers using the programs Excel (Microsoft), Word 5.1 (Microsoft), Illustrator v5.0 (Adobe) and Cricket Graph III.

3.1 Introduction

Transgenic animals are extensively used to analyse hypermutation. Transgenes containing relevant immunoglobulin sequences mutate (O'Brien et al., 1987) and the rate of hypermutation is of the same order of magnitude as the endogenous genes when the elements in the 3' flank of the κ locus are included (Sharpe et al., 1991). So far this is the assay system most widely used to analyse the regulatory elements controlling hypermutation.

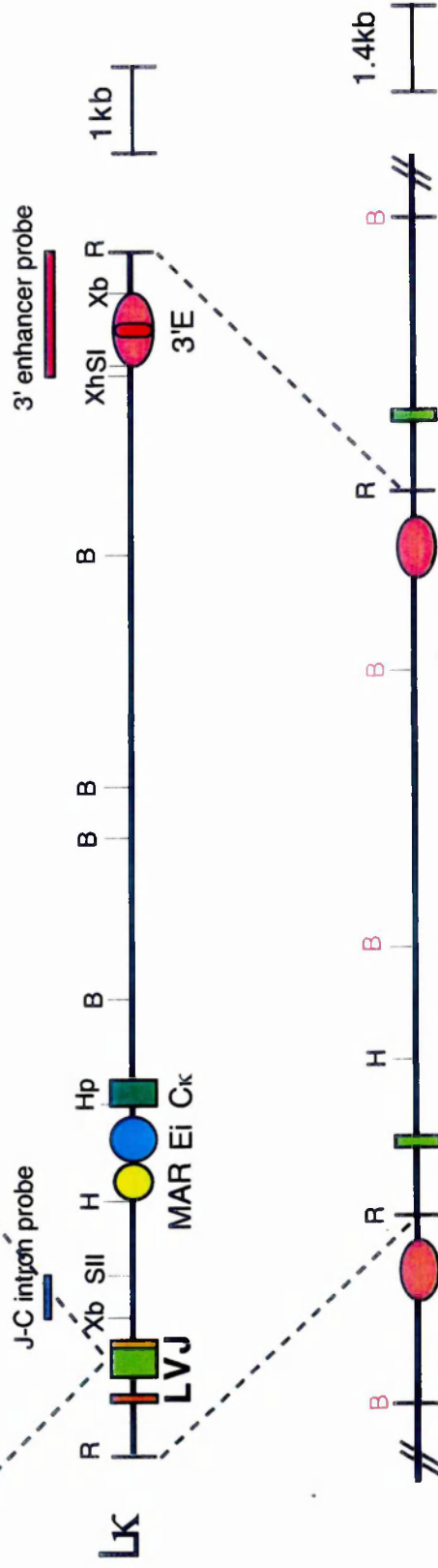
Transgenic constructs containing modified versions of the relevant κ chain transgene have been used to identify putative regulatory sequences essential for hypermutation (Betz et al., 1994; Yélamos et al., 1995). These studies have shown that the 3' enhancer and κ intron enhancer/matrix attachment region are crucial to obtain high levels of mutation and that the lymphoid-specific promoter elements and the V κ gene fragment are not essential to recruit hypermutation. However, the essential regions (3'E and Ei/MAR) have not been further dissected to obtain a tighter definition of the *cis*-acting DNA sequences required to target somatic mutation.

Transgenic mice also have been valuable in the study of the intrinsic properties of the hypermutation machinery. The analysis of somatic mutation in passenger transgenes makes it possible to scan all positions within the V gene for inherent mutability, since they do not contribute to the antigen specificity of the B cell (Betz et al., 1993b; Sharpe et al., 1991). These studies showed that the mutations that are not distributed randomly along the V gene. Instead certain positions are favoured sites for mutations (hotspots) while others are rarely mutated (coldspots).

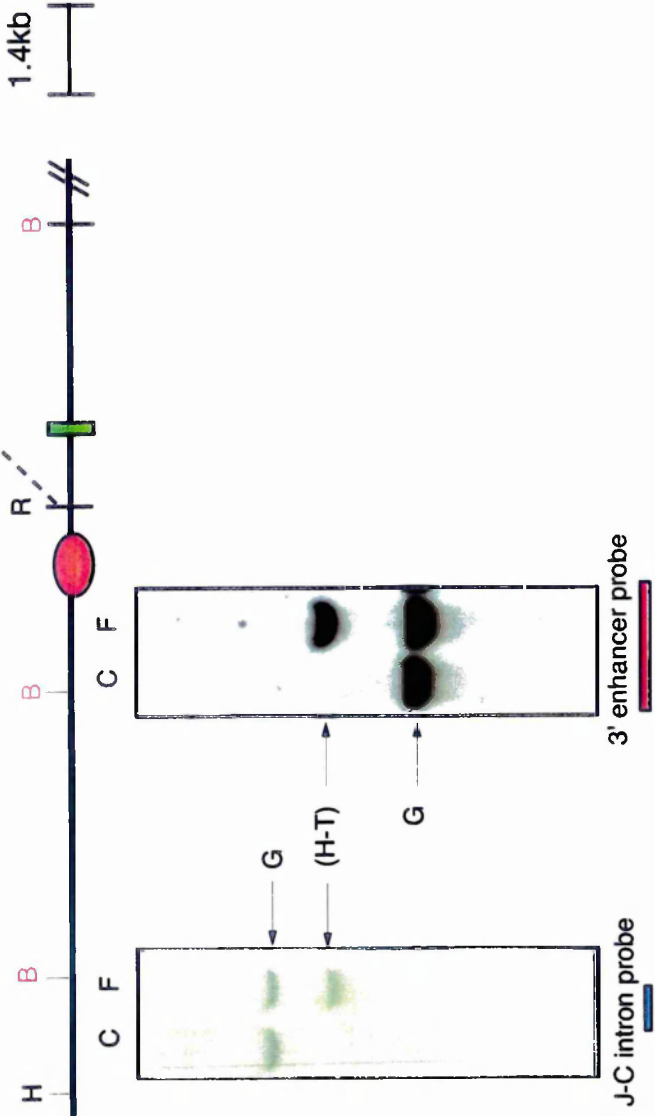
Mutational hotspots in immunoglobulin genes are only in part created by their DNA sequence since they are not found in all positions conforming to the hotspot consensus. The contribution of the local DNA secondary structure and the primary sequence of and close to the hotspots remains unresolved.

In the absence of an *in vitro* system, transgenic mice which carry modified rearranged immunoglobulin genes represent the only system currently available to identify the elements that regulate and target hypermutation. Several lines of transgenic mice were generated to analyse which DNA segments within the κ 3' enhancer and MAR/Ei regions are required to target somatic mutation and how the local DNA secondary structure and the primary sequence effect hotspots. The description of the constructs, used to generate those lines, can be found in Materials and Methods.

a GC**G** AGC TC**G** AGC GTT AGC TAT ATG CAT **Lk7M**



b



3.2 Transgenic lines to study the role of the DNA sequence and secondary structure in hypermutation: Lκ(WT&7M), LκNG and Lκ7M/TCA&Δ[Li]

Three transgenic mouse lines were produced to examine whether the DNA secondary structure and the primary sequence are important for hypermutation. The lines contain synthetic Vκ domains with modifications in the secondary structure and the consensus sequences in the CDR1. This region was chosen because it is usually more strongly mutated than other regions (Malipiero et al., 1987; Betz et al., 1993b). It contains a large accumulation of mutational hotspots including the residue Ser31, which is the hotspot most frequently identified in some of the V genes analysed. This Ser31 is encoded by a AGT triplet, which falls within the hypermutation hotspot consensus Pu-G-Py-A/T.

The transgenic lines were generated using a non conventional approach. An equimolar mixture of different transgene fragments was injected in the same mouse zygote. The transgenic mice created in this way contain multiple, distinctly modified transgene copies. Several transgenes could therefore be tested in the same animal and can served as internal control for each other.

3.2.1 Lκ(WT&7M) line

An equimolar mixture of eight different DNA fragments was used for microinjection of mouse eggs. Seven of them (7M, M25, M28, M31, M30, M36 and M52) carry several silent distinguishable mutations in the CDR1 (see Appendix I B). The eighth was the control LκWT. These mutations were designed to disrupt the consensus sequence and the secondary structure in this region.

One transgenic founder was obtained. The number of copies of the integrated DNA was estimated by Southern blot using tail DNA digested with BamHI. The blot was hybridised first with the XbaI-SacII fragment of the J-C intron and then rehybridised with the XhoI-EcoRI fragment that includes the 3' enhancer region (Fig. 3.1). The two endogenous gene copies in the germline configuration allowed the estimation of transgene copy number.

Figure 3.1 Genetic map of transgenes in Lκ(WT&7M) line (opposite page).

a) Diagram of LκWT and Lκ7M transgenes. The Lκ7M encodes a Vκ region containing 7 silent mutations (denoted in red) in the CDR1. DNA fragments used as probes are indicated. Restriction sites are abbreviated: R, EcoRI; B, BamHI; Xb, XbaI; Xh, XhoI; SI, SacI; SII, SacII; H, HindIII.

b) The genetic map deduced by Southern blot analysis of tail DNA from the Lκ(WT&7M) founder (F) and from a non transgenic litter mate (C). The DNA was digested with BamHI. The blot was hybridised first with the J-C intron probe and then rehybridised with the 3' enhancer probe. The 8.9kb fragment derived from head-to-tail (H-T) junctions. The germline configuration (G) is indicated.

Southern blot analysis disclosed that in addition to the fragment which originates from the germline there is another fragment that originates from the head-to-tail (H-T) (the head corresponds to the 5' end and the tail corresponds to the 3' end of the constructs) junction of the transgenes. The transgene is seen as a 8.9kb fragment after hybridisation with both, the J-C intron and the 3'E probes. The intensity of the band corresponds to two H-T junctions (Fig. 3.1).

Southern blots using tail DNA digested with several restriction enzymes confirmed that there were three copies of the transgene in the two H-T junctions. Two of the copies are incomplete, one lacking the 5' region and the other lacking the 3' region of the transgene (Fig. 3.1).

The sequence analysis of hybridomas revealed that the two V κ segments present originate from the L κ WT and the L κ 7M. Due to the particular symmetry of the tandem array it was not possible to determine which of the two was the complete copy.

Comparison of the tail blots of the founder with those of its transgenic offspring showed that the transmission of the introduced DNA followed simple Mendelian rules.

3.2.2 L κ NG line

An equimolar mixture of six different vector-free DNA fragments was used to inject mouse zygotes. Fragments were obtained from L κ Δ B7M, L κ Δ BTCA (M60), L κ Δ B7M/TCA (M72), L κ -Vgpt*, L κ -Vneo* [XS]i and the control L κ Δ BWT. The first three constructs carry several silent distinguishable mutations in the CDR1 (see Appendix I B).

The constructs L κ -Vgpt* and L κ -Vneo* [XS]i were designed to determine whether non Ig sequences can be a target for hypermutation (Yélamos et al., 1995). The analysis of these two constructs was not part of my work and is not included in this thesis. This transgenic line is a good example how different transgenes can be analysed in a single mouse, and how one set of constructs can provide an internal control for another.

One founder was obtained. Southern blot analysis revealed that it carries several copies, of which some are in H-H, T-T or T-H configuration (data not shown). In addition there were several copies in head to host DNA arrangement. This suggests that not all of them belong to a unique multi-copy array. The number of copies estimated by the Southern blot revealed with the J-C intron probe (about 10) does not correspond with the number estimated when the blot was revealed with the 3' enhancer probe (about 6). This indicates that several copies are truncated.

The sequence analysis of hybridomas shows that in addition to the L κ -Vgpt* and L κ -Vneo* [XS]i two L κ Δ B7M and three L κ Δ BWT copies are present. One of the WT copies carries a germline point mutation (G for A) introduced fortuitously at position 54 of the V region.

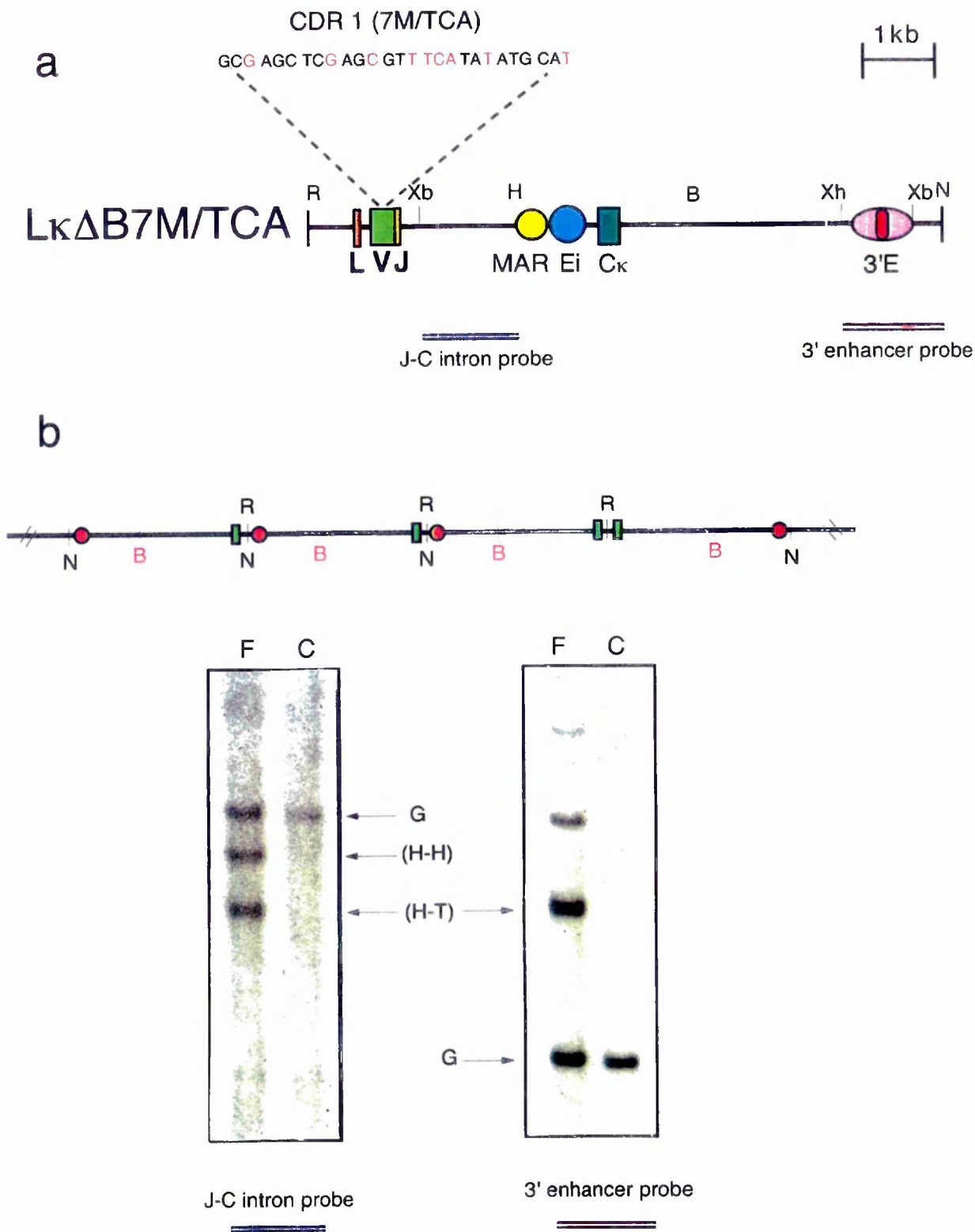


Figure 3.2 Schematic representation of L κ Δ B7M/TCA transgene and genetic map of transgenes in L κ 7M/TCA&[Li] line. a) Diagram of L κ Δ B7M/TCA transgene. The silent mutations introduced in the CDR1 are shown in red. DNA fragments used as probes are indicated. Restriction sites are abbreviated: R, EcoRI; B, BamHI; Xb, XbaI; Xh, XhoI; N, NotI; H, HindIII. b) The genetic map deduced by Southern blot analysis of tails DNA from L κ 7M/TCA&[Li] founder (F) and a non transgenic litter mate (C). The DNA was digested with BamHI. The blot was hybridised first with the J-C intron probe and then rehybridised with the 3' enhancer probe. The 8.9kb fragment is derived from head-to-tail (H-T) and the 10.6kb fragment derived from head-to-head (H-H) configurations. The additional two bands are thought to be T-host genome fragments. The germline fragment (G) is also indicated.

3.2.3 Lκ7M/TCA&Δ[Li] line

The founder of this transgenic line was obtained by microinjection with the equimolar mixture of the vector free fragments obtained from the LκΔB7M/TCA (Fig. 3.2), which carries 9 silent mutations in the CDR1, and the LκΔ[Li] constructs. The LκΔ[Li] is an Lκ derivative carrying a deletion in the 5' leader intron (C. Rada, unpublished). Its analysis is not included in this thesis.

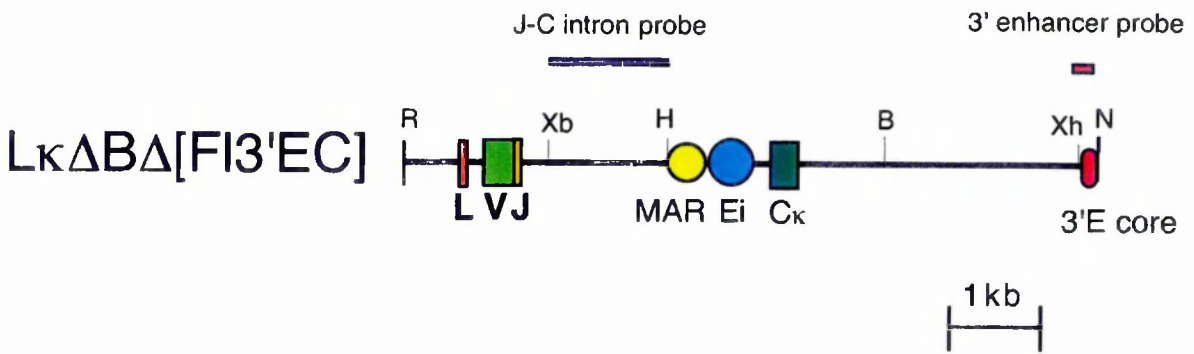
Southern blot analysis of the founder and its offspring disclosed that in addition to the germline V κ fragment there is a fragment that corresponds to the head-to-tail (H-T) configuration of the transgene. The fragment is seen as a 8.9kb band after hybridisation with both, the J-C intron and the 3'E probe. The intensity of the band corresponds to two H-T junctions (Fig. 3.2). There is also a 10.6kb band after hybridisation with the J-C intron probe that corresponds to one H-H junction. Two additional bands are disclosed after hybridisation with the 3' enhancer probe. They are big enough to include the complete BamHI fragment which contains the 3' enhancer and are therefore probably tail towards host genome fragments.

These results suggest that there are at least 4 complete copies of the transgene. Sequence analysis of hybridomas disclosed at least one copy of LκΔB7M/TCA transgene in addition to LκΔ[Li]. The genetic map deduced from the analysis of the fragments obtained in the Southern blot is shown in Fig 3.2b. In the generations inspected so far the transgenes were transmitted with the predicted Mendelian frequency.

Although the distance between the V and the leader region is different in the two transgenes of this line, it is not big enough for discrimination by Southern blots. However both transgenes could easily be distinguished by appropriate PCR amplification (see Chapter 4). They could also be identified due to specific restriction sites introduced in the CDR1 of Lκ7M/TCA. In the sequence analysis, both V regions could be recognised by the mutations specific for Lκ7M/TCA.

All these factors are important for the design of mixtures with different constructs to enable the analysis of the independent transgenes.

a



b

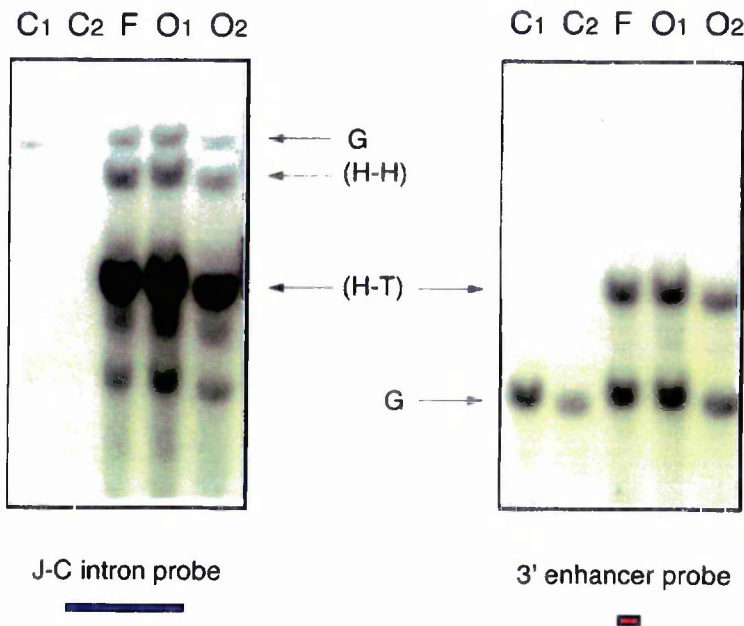


Figure 3.3 Schematic representation of $L_{\kappa}\Delta B\Delta[F13'EC]$ transgene and Southern blot analysis of $L_{\kappa}\Delta[F13'EC]$ line. a) Diagram of $L_{\kappa}\Delta B\Delta[F13'EC]$ transgene. Fragments used as probes are shown. Restriction sites are abbreviated: R, EcoRI; B, BamHI; Xb, XbaI; Xh, XhoI; N, NotI; H, HindIII. b) Southern blot analysis of tails DNA from $L_{\kappa}\Delta[F13'EC]$ founder (F), two non transgenic litter mates (C), and two of the founder offspring (O1 and O2). The DNA was digested with BamHI. The blot was hybridised first with the J-C intron probe and then rehybridised with the 3' enhancer probe. The 7.7kb fragment is derived from the head-to-tail (H-T) and the 10.6kb fragment derived from the head-to-head (H-H) configurations. The germline fragment (G) is also indicated. The additional bands are thought to be H-host genome integration.

3.3 Transgenic lines to study the role of the 3' enhancer core and its flanking sequences: LκΔ[3'EC] and LκΔ[F13'EC]

3.3.1 LκΔ[F13'EC] line

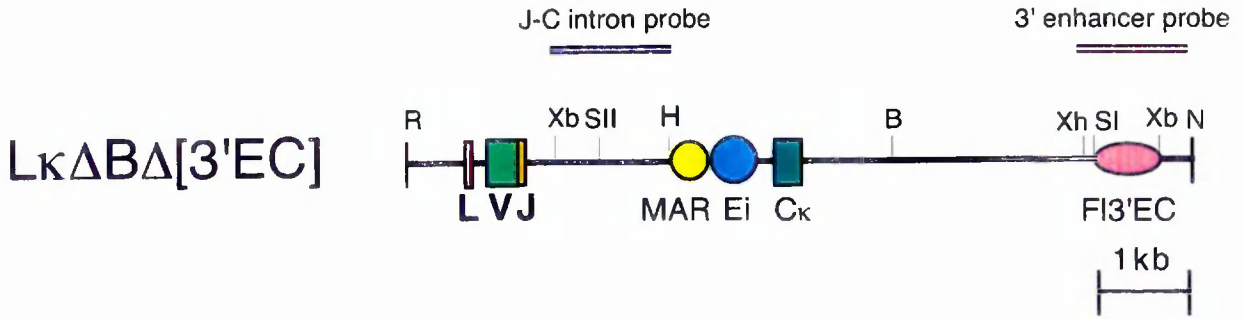
This line was made by pronuclear microinjection with the vector free fragment of LκΔBΔ[F13'EC] (Fig. 3.3), which carries the deletion of the 3' enhancer core flanking regions. Tail DNA digested with BamHI was used in Southern blots and hybridised with the J-C intron probe and the XhoI-NotI fragment containing the 3' enhancer core region. The approximate copy number estimation was done as described in the previous sections, by comparing the intensity of the fragments with the band obtained for the two endogenous copies in germline configuration.

There is a 7.7kb band after hybridisation with both the J-C intron and 3'E probe, which corresponds to a head-to-tail (H-T) configuration (Fig. 3.3). The fragment is seen with the same intensity of the germline band when probed with the 3' fragment. However it is 4-6 fold stronger when the J-C intron probe was used. This difference may be due to the presence of other fragments which contain only the J-C intron region and migrate very close to the H-T band.

There is also a 10.6kb band after hybridisation with the J-C intron probe, which corresponds to two H-H junctions. Two additional bands are disclosed after hybridisation with the J-C intron probe. They are big enough to include the complete BamHI fragment of the VK region and therefore they may be head oriented towards host genome fragments.

The presence of many fragments (at least 10) after the J-C intron probe is used and the absence of the corresponding amount of fragments after probing with the 3' region, suggests the presence of many incomplete transgenes which contain the 5' region but lack their 3' region of the construct. These results suggest that there are only two complete copies of the transgene.

a



b

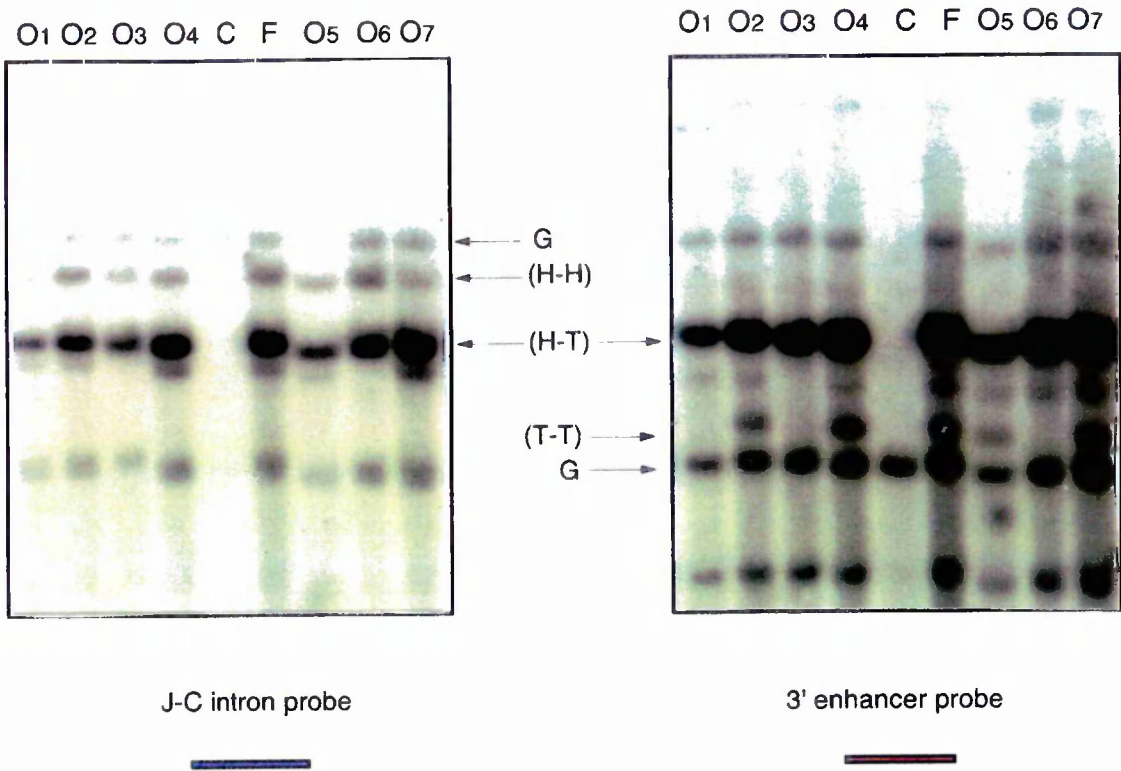


Figure 3.4 Schematic representation of LkΔBΔ[3'EC] transgene and Southern blot analysis of LkΔ[3'EC] line. a) Diagram of LkΔBΔ[3'EC] transgene. Fragments used as probes are shown. Restriction sites are abbreviated: R, EcoRI; B, BamHI; Xb, XbaI; Xh, XhoI; SI, SacI; N, NotI; H, HindIII. b) Southern blot analysis of tails DNA from the LkΔ[3'EC] founder (F), a non transgenic litter mate (C) and seven of the transgenic offspring (O(1-7)). The DNA was digested with BamHI. The blot was hybridised first with the J-C intron probe and then rehybridised with the 3' enhancer probe. The 8.7kb fragment is derived from the head-to-tail (H-T), the 10.6kb fragment is derived from the head-to-head (H-H), and the 6.8kb fragment derived from the tail-to-tail (T-T) configurations. The germline fragment (G) is also indicated. The additional bands are thought to be tail and head towards host genome integration.

3.3.2 L κ Δ [3'EC] line

This line was made by pronuclear microinjection with the vector free fragment of L κ Δ B Δ [3'EC] (Fig. 3.4), which carries the deletion of the core enhancer. Southern blots of tail DNA digested with BamHI were hybridised with the J-C intron probe and the XhoI-Not fragment containing the 3' enhancer flanking regions.

There is a 8.7kb band after hybridisation with the J-C intron and 3'E probes which corresponds to the head-to-tail (H-T) configuration (Fig. 3.4). This fragment is 4 times more intense than the endogenous band. There is also a 10.6kb band after hybridisation with the J-C intron probe that corresponds to a H-H junction. In addition a 6.8kb band is disclosed when the 3'E probe is used that corresponds to a T-T junction. Additional fragments are thought to be transgene-host genome junctions (four tail-to-mouse and two head-to-mouse) suggesting that the transgenes are in different multy-copy arrays. These results suggest that the transgenic L κ Δ [3'EC] founder carries at least ten complete copies of the transgene.

Comparison of the tail blots from the founder with those from its transgenic offspring shows some differences. The fragment which corresponds to the tail-to-tail junction was missing in 20% of the transgenic mice analysed. One of the four fragments thought to be tail-to-host genome integration was lost in 13% of the transgenic mice and both fragments were missing in 10% of the transgenic offspring. Therefore, 43% of the transgenic mice in this line lost at least one fragment. One of the bands thought to be head-to-host genome integration was also lost in some of the offspring.

The copy loss was corroborated by breeding experiments. A founder descendent containing all the copies was crossed with F1 (CBAxC57BL/6J). 50% of the transgenic offspring contained deletions. In addition, each deletion variant was crossed with F1 mice and the progeny analysed by Southern blotting. The results show that the progeny retains the expected parental phenotype. The number of mice analysed however is so far not large enough to conclude that all the deletion variants are stable.

The instability of long inverted repeat structures (such as head-to-head and tail-to-tail), which occurs upon transgene integration, has also been observed in other transgenes (Collick et al., 1996). The mutation process involved can lead to both insertions or deletions resulting in the stabilisation of the transgene.

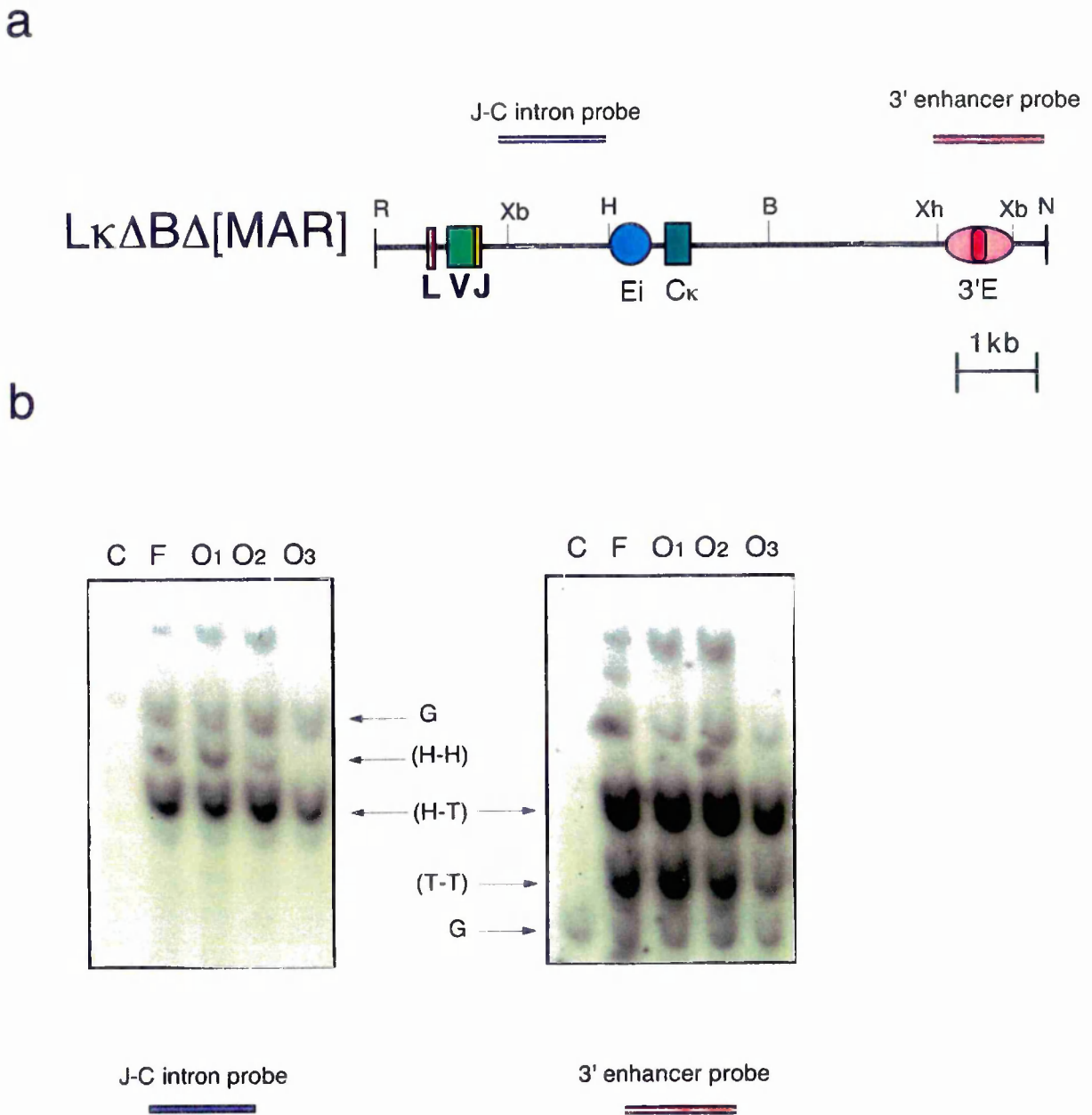


Figure 3.5 Schematic representation of $L\kappa\Delta B\Delta[MAR]$ transgene and Southern blot analysis of $L\kappa\Delta[MAR]$ line. a) Diagram of $L\kappa\Delta B\Delta[MAR]$ transgene. The fragments used as probes are shown. Restriction sites are abbreviated: R, EcoRI; B, BamHI; Xb, XbaI; Xh, XhoI; N, NotI; H, HindIII. b) Southern blot analysis of tails DNA from the $L\kappa\Delta[MAR]$ founder (F), a non transgenic litter mate (C) and three of the transgenic offspring (O(1-3)). DNA was digested with BamHI. The blot was hybridised first with the J-C intron probe and then rehybridised with the 3' enhancer probe. The 8.6kb fragment is derived from the head-to-tail (H-T) integration, the 10kb fragment is derived from the head-to-head (H-H) and 7.2kb fragment is derived from the tail-to-tail (T-T) configurations. The germline fragment (G) is also indicated. The additional bands are thought to be tail and head towards host genome integrations.

3.4 Transgenic lines to study the role of the DNA elements within the J-C intron: LκΔ[MAR] and LκΔ[Ei]

3.4.1 LκΔ[MAR] line

This line was made by microinjection with the vector free fragment of LκΔBΔ[MAR] (Fig. 3.5), which carries the deletion of the MAR.

Tail DNA digested with BamHI was used in Southern blots and hybridised with the J-C intron and 3' enhancer probes. Southern blot analysis revealed that the LκΔ[MAR] founder has a 8.6kb band after hybridisation with both probes which corresponds to the head-to-tail (H-T) configuration (Fig. 3.5). This fragment is at least 4 times more intense than the endogenous band. There is also a 10kb band after hybridisation with the J-C intron probe that corresponds to a H-H junction. In addition a 7.2kb band is disclosed when the 3'E probe is used which corresponds to two T-T junctions. Additional fragments of different size are probably transgene-host genome junctions (three tail-to-mouse and two head-to-mouse) suggesting that the transgenes are in different multi-copy arrays. These results suggest that there are at least 11 complete copies of the transgene. The transmission of the transgene was studied by Southern blotting of the founder offspring. It was found that 15% of the transgenic mice had deleted three fragments: the fragment diagnostic for the head-to-head integration, one of the tail to mouse and one of the head-to-mouse fragments (see offspring O₃ in Figure 3.5). These results demonstrate the instability of head-to-head integrations, possibly originated by the inverted repeat.

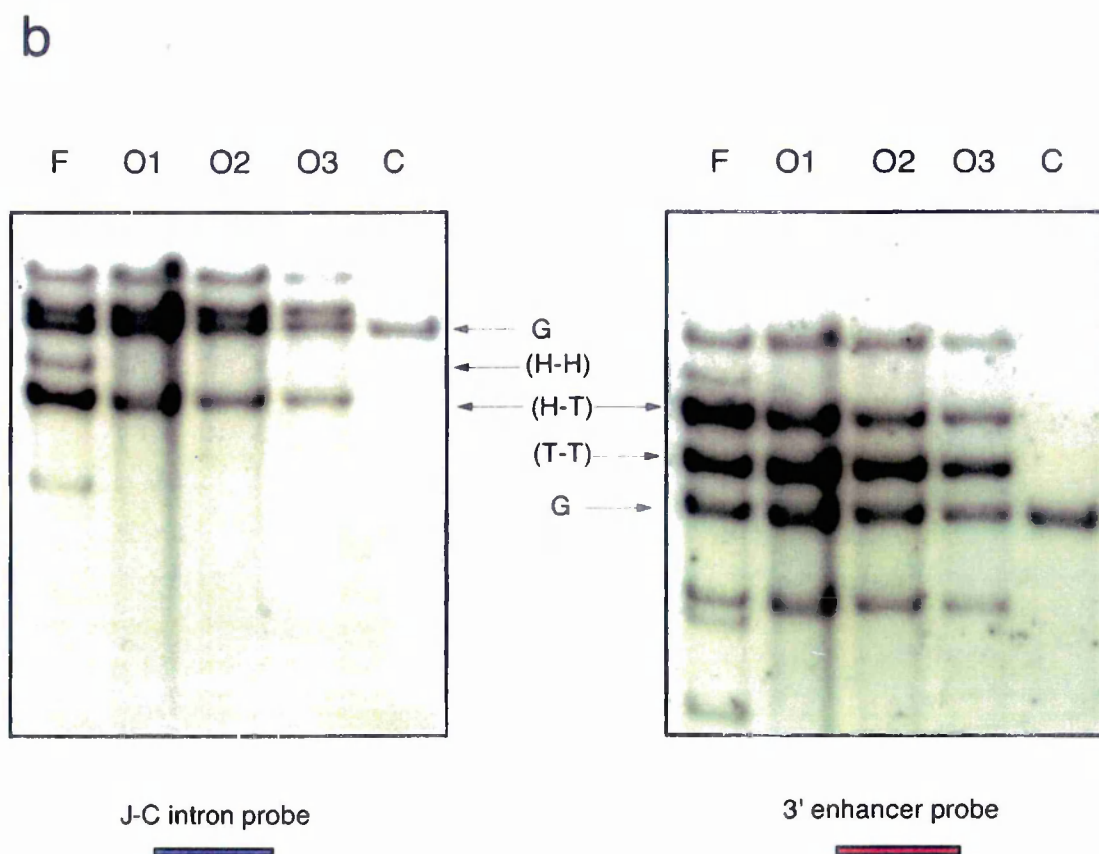
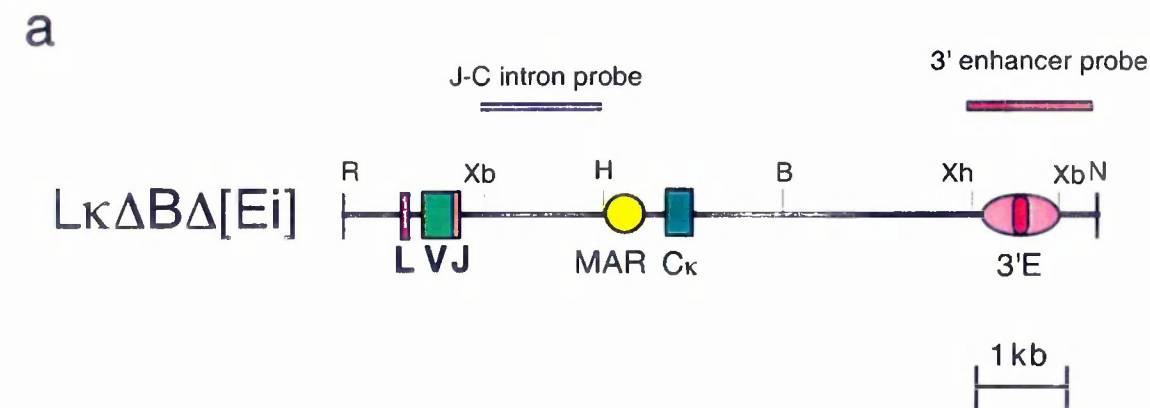


Figure 3.6 Schematic representation of $L\kappa\Delta B\Delta[Ei]$ transgene and Southern blot analysis of $L\kappa\Delta[Ei]$ line a) Diagram of $L\kappa\Delta B\Delta[Ei]$ transgene. The fragments used as probes are shown. Restriction sites are abbreviated: R, EcoRI; B, BamHI; Xb, XbaI; Xh, XhoI; N, NotI; H, HindIII. b) Southern blot analysis of tails DNA from $L\kappa\Delta[Ei]$ founder (F), a non transgenic litter mate (C) and three of the transgenic offspring (O(1-3)). The DNA was digested with BamHI. The blot was hybridised first with the J-C intron probe and then rehybridised with the 3' enhancer probe. The 8.5kb fragment is derived from the head-to-tail (H-T), the 9.8kb fragment is derived from the head-to-head (H-H) and the 7.2kb fragment is derived from the tail-to-tail configurations. The germline fragment (G) is also indicated. The additional bands are thought to be tail and head toward host genome integrations.

3.4.2 L κ Δ [Ei] line

This line was obtained by microinjection of the mouse zygote with the vector free L κ Δ B Δ [Ei] fragment (Fig. 3.6), which carries the deletion of the κ intron enhancer region.

Southern blot analysis of tail DNA digested with BamHI revealed that the L κ Δ [Ei] founder has a 8.5kb band after hybridisation with the J-C intron and 3'E probes, which correspond to the head-to-tail (H-T) configuration (Fig. 3.5). This fragment is twice as intense as the endogenous band. There is also a 9.8kb band after hybridisation with the J-C intron probe which could correspond to a single H-H junction, although it is slightly weaker than the endogenous band. It may be that part of the J-C intron fragment used as a probe has been lost as a consequence of a rearrangement with the host genome. In addition a 7.2kb band is disclosed when the 3'E probe is used that corresponds to two T-T junctions. Additional fragments of various size are presumably transgene-host genome junctions (six tail-to-mouse and three head-to-mouse) suggesting that the transgenes are in different multi-copy arrays. These results suggest that there are at least 8 complete copies of the transgene.

The transgene transmission was studied by Southern blotting of the founder offspring. It was found that all the transgenic progeny had deleted the fragment in the head-to-head configuration, one of the head-to-host and three tail-to-host fragments. In addition the band corresponding to T-H configuration was weaker (it has the same intensity as the endogenous gene in the germline configuration) than in the founder. These results further confirm the instability of the inverted repeat structures.

3.5. Expression analysis

The level of expression of the transgenes is usually determined in the serum of the transgenic mice, in the NS0 cells stably transfected or in hybridomas derived from the transgenic line. The analysis of expression of immunoglobulin transgenes is complicated by a number of interrelated phenomena peculiar to the immune system.

Several reasons could disguise the real effect of the modifications that the transgene carries on the expression. The expression can be affected by the chromatin status in the site of the transgene integration, by the number and type of cells that are expressing the transgene and or by the efficiency of the transgene in mediating allelic exclusion. Also, the chromosomal rearrangements that occurs in fusions can up or down regulate the immunoglobulin gene expression of hybridoma cells. The age of the mice when the serum is collected is also important. In the transfected NS0 cells the SV40 enhancer present in the vector can promote expression of the immunoglobulin gene. Therefore, the quantitative results derived from these systems should not be taken as an absolute measure of the effects that the constructs incur.

3.5.1 Transgene κ chain expression in hybridomas derived from the transgenic lines

Hybridomas were produced for all the transgenic lines obtained (except for L $\kappa\Delta$ [Ei]). Mice were immunised and boosted with pHx-CSA (see Materials and Methods). The proportion of hybridomas expressing only the transgenic rat κ chain, the endogenous mouse κ or both the endogenous mouse and the transgenic rat κ chains were determined. This gives an indication whether or not the level of transgene expression are effective in mediating allelic exclusion of the endogenous κ gene.

The L $\kappa\Delta$ B transgene allowed effective feedback inhibition of the endogenous κ chain in all the hybridomas analysed derived from L κ NG transgenic line. Similar result (80-90%) was found in hybridomas obtained from the L κ 6 line, which carries the L κ construct (Meyer et al., 1990).

However, only 36% of the hybridomas obtained from the L κ (WT&7M) mediated allelic exclusion. Hybridomas derived from L $\kappa\Delta$ [F13'EC] line were also relatively ineffective in mediating allelic exclusion (31%). These results may be a consequence of the low number of intact copies found in those lines (1 and 2 respectively).

In hybridomas derived from L $\kappa\Delta$ [3'EC] and L κ 7M/TCA& Δ [Li] a relatively good allelic exclusion (76 and 60% respectively) was observed.

3.5.2 Transgene κ chain expression in the serum of the transgenic mice

As a first step towards quantitation of transgene expression the levels of rat Ig κ in the serum of the transgenic mice were analysed by ELISA calibrated using P4C1/13-11 (fraction I) antibody. In this inhibition assay the test serum is incubated with an anti-rat κ reagent and the remaining free antibody is quantitated. Dilutions of test serum were incubated o.n. with a mouse anti-rat κ antibody (MARK-1). The mixture was then transferred to the plate coated with rat antibody (P4C1/13-11, fraction I). The amount of MARK-1 Ab bound was assayed by incubation with biotinylated sheep anti-mouse antiserum, which was detected with a Streptavidin-HRP conjugate. Samples were taken from mice at the age of approximately 4 weeks.

The results of the ELISA performed with the founder sera are shown in Fig. 3.7a. This data originates from one assay with duplicate samples. These results have been repeated and confirmed (data not shown). The serum of five offspring was pooled and analysed several times, always in duplicate. The results are shown in Fig. 3.7b. The standard deviation is indicated.

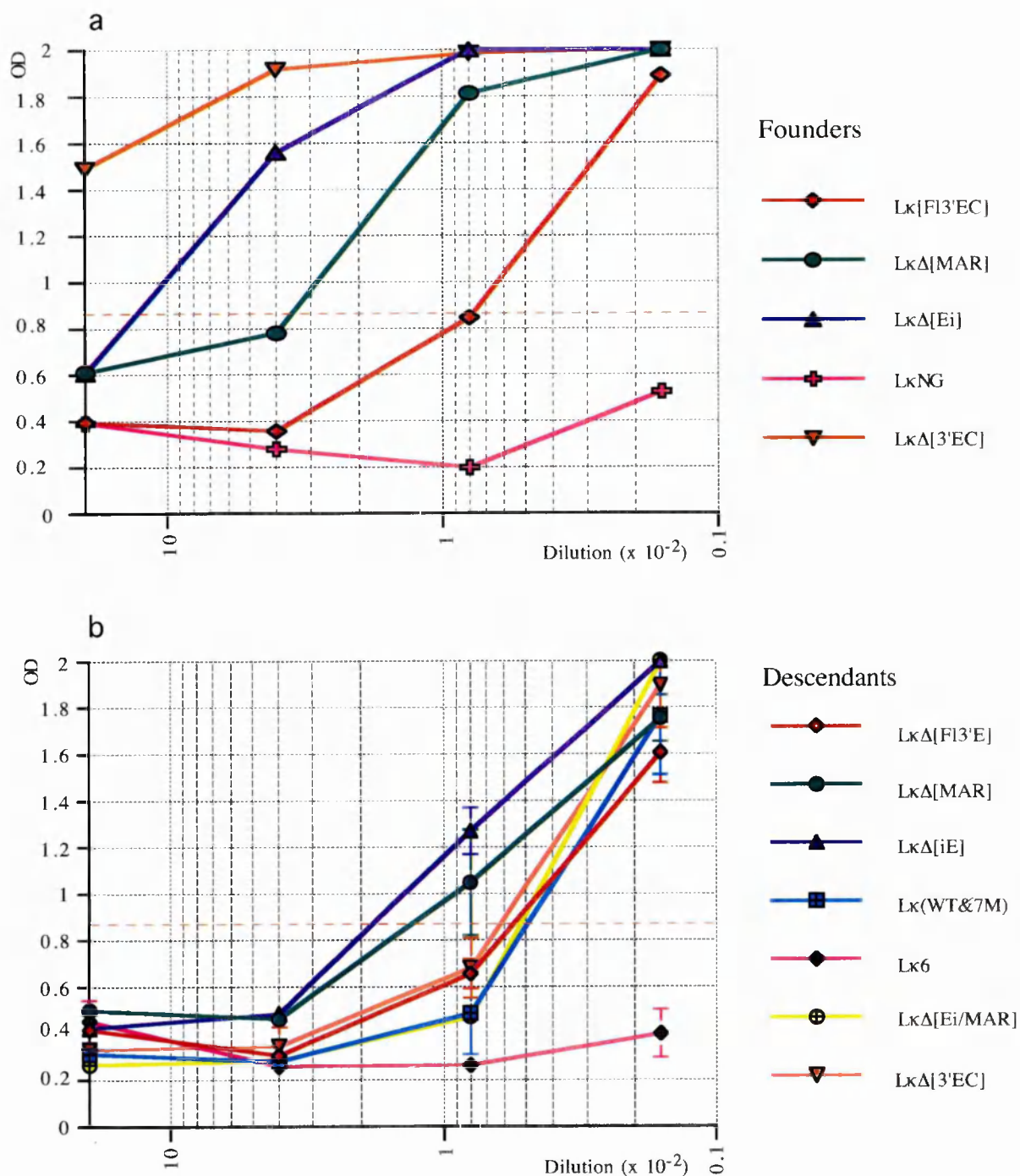


Figure 3.7 Transgene expression in the mouse sera for each of the transgenic lines. The sera (diluted 1/5; 1/25; 1/125 and 1/625) was tested for κ rat expression by ELISA. Serial dilutions of known concentrations of P4C1($\mu\text{g/ml}$) were used as standard curve. The broken pink horizontal line shows the OD corresponding to 50% inhibition. The equivalent concentration for this OD value is 2 $\mu\text{g/ml}$.

a) Founders.

b) Pooled offspring. Values represent the mean of several assays. The standard deviation is indicated. The results from Lκ6 and LκΔ[Ei/MAR] mice are shown for comparison.

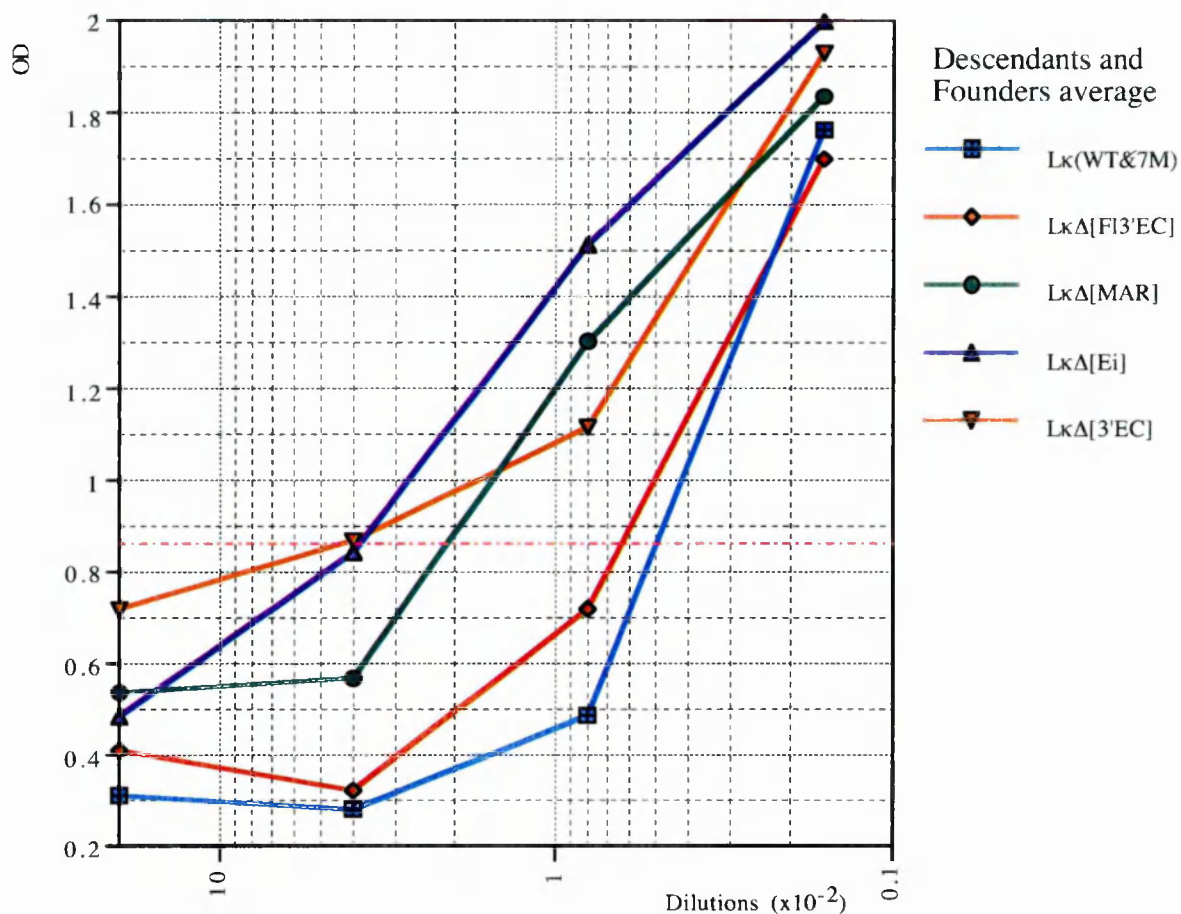


Figure 3.8 Average transgene expression in the mouse sera for each of the transgenic lines.

The average of the transgene expression in the mouse sera (diluted 1/5; 1/25; 1/125 and 1/625) of the founder and their offspring is shown. The broken pink horizontal line shows the OD. corresponding to 50% inhibition (0.85). The equivalent concentration for this OD is 2 μ g/ml.

The analysis of the κ rat expression revealed that the titre of serum antibody differs between the founder and its descendants in some transgenic lines. In L κ Δ [MAR] and L κ Δ [Ei] the offspring expressed only slightly more (2.5 and 7 times respectively) but in the case of L κ Δ [3'EC] line the difference is more significant (30 times). The detailed reasons for this phenomenon are unknown, but it is important to note that these are the same transgenic lines, for which the copy loss was observed. It is possible that the expression of the transgene has been upregulated as a consequence of the stabilisation process. Fig. 3.8 shows the average pooled titres of the serum of founders and its descendants.

The transgene is expressed in all the lines analysed, but only in the L κ NG line - which carries 5 copies of the L κ Δ B transgene - expression is comparable to that of the L κ 6 line, which carries 5 copies of L κ . The estimated titre of transgenic κ chains in the L κ 6 mice is higher than 1mg/ml (extrapolated in Fig. 3.7b) which agrees with the previously determined values of 1-4 mg/ml (Meyer et al., 1990).

For the L κ (WT&7M) line, which contains only two V κ regions (one belongs to an intact copy; the other is truncated, while having the 3'E in a 5' position with respect to the V), the estimated titre is 400 μ g/ml. This value is 2.5-10 times lower than that of the L κ NG and L κ 6 lines.

In the transgenic line carrying the deletion of the 3' enhancer core flanking regions the expression of the transgenic κ chain was only reduced about 3 to 13 times with respect to the values observed in L κ NG and L κ 6 lines. The level of transgene expression in the founder of the line containing the deletion of the core 3' enhancer was significantly reduced (250-1000 times lower) than that of the L κ NG and L κ 6 lines.

This large reduction in the transgene expression however was not observed when the pooled serum from the founder's offspring was assayed, in which it was only about 10 times lower than in the L κ NG and L κ 6 lines. As I explained before, the transgene expression in the offspring may be upregulated, therefore it is essential to obtain and analyse more founders to confirm these results.

The level of transgene κ light chain expression in the lines containing deletions of MAR and Ei, are lower than that of L κ NG and L κ 6 lines. The values are 19-77 and 71-286 times lower respectively for the founders and 7-30 and 10-40 times lower respectively for the descendants.

These results indicate that all the regulatory elements analysed, particularly the core of the 3' enhancer, are required to achieve high levels of κ light chain expression.

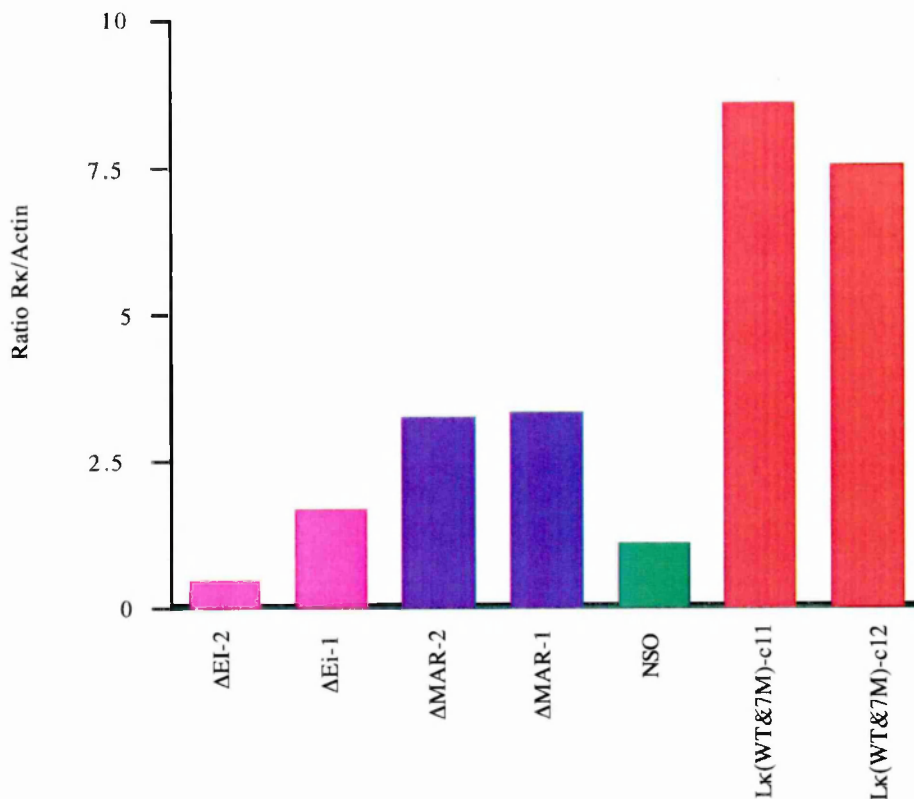


Figure 3.9 Expression of the LκΔ[Ei] and LκΔ[MAR] transgene in NSO stable transfectants.

Northern blot assay of RNA from the transfected κ gene of NSO cells stably transfected with LκΔ[Ei] and LκΔ[MAR] constructs. For each construct transcripts of the transfected κ gene and the endogenous β-actin were assayed in samples of RNA prepared from pools of 10 wells each of which contained multiple clones. Quantitation by densitometric analysis were made and the values for rat are plotted relative to actin. LκΔ[Ei] and LκΔ[MAR] are NSO transfectants and Lκ(WT&7M) are hybridomas.

3.5.3 Transgene κ chain expression in stably transfected NS0 cells

The transgene expression in supernatants from NS0 cells stably transfected with the constructs L κ Δ [Ei] and L κ Δ [MAR] were also determined. Although quantitative analyses were not done, "+" signs were given as indication of the relative O.D. in the assays. The level of expression found in the L κ Δ [Ei] transfectants was very low (+) (it was only detectable in the neat supernatant) on the contrary the L κ Δ [MAR] transfectants expressed well (+++).

The transgene expression in NS0 cells stably transfected with L κ Δ [Ei] and L κ Δ [MAR] transgenes was also analysed on the RNA level by Northern blots. Quantitation by densitometric analysis of the transgenic κ chain were performed and the values were corrected by the corresponding actin values (Fig. 3.9). The results for NS0 cells and two L κ (WT&7M) hybridomas are included for comparison. Two independent transfectant pools were tested for each transgene. The results indicate that the expression levels in the case of the transfectants carrying the deletion of the Ei are hardly significant and lower than after deletion of MAR. Both are lower than in the L κ hybridoma.

These results further confirm that the Ei and MAR elements are required for high level of expression. The characteristics of all the transgenic lines described in this study are summarised in Table 3.1.

Table 3.1 Summary of the characteristics of the transgenic lines .

Transgenic line	DNA copy # number	Transgene instability \approx	Transgene expression *			Allelic exclusion in hybridomas
			mouse sera Titre		stable [¶] transfectants	
			(F)	(D)		
L κ (WT&7M)	1	NF	ND	400	(++++)	36%
L κ NG	10	ND	>1000	ND	(++++)	100%
L κ 7M/TCA& Δ [Li]	4	NF	ND	ND	(++++)	60%
L κ Δ [F13'EC]	2	NF	250	333	ND	31%
L κ Δ [3'EC]	10	+	<4	308	ND	76%
L κ Δ [MAR]	11	+	52	133	(+++)	ND
L κ Δ [Ei]	8	+	14	100	(+/-)	ND
L κ 6 [§]	5	ND	-	1000-4000	-	90%

Approximate number of complete copies of the transgene.

\approx Copy loss is indicated "+". ND stands for not determined. NF stands for not found.

* Titres are given as the concentration of the transgene in the mouse sera (μ g/ml) required to sequester 50% of the anti-rat antibody used in the ELISA in the serum of founders (F) and their descendants (D).

[¶] Transgene expression in the supernatant of the NS0 cells stable transfected was tested by ELISA, the number of crosses refers to the relative O.D. in the assays.

[§] Data from L κ 6 line (Meyer et al., 1990) are provided for comparison.

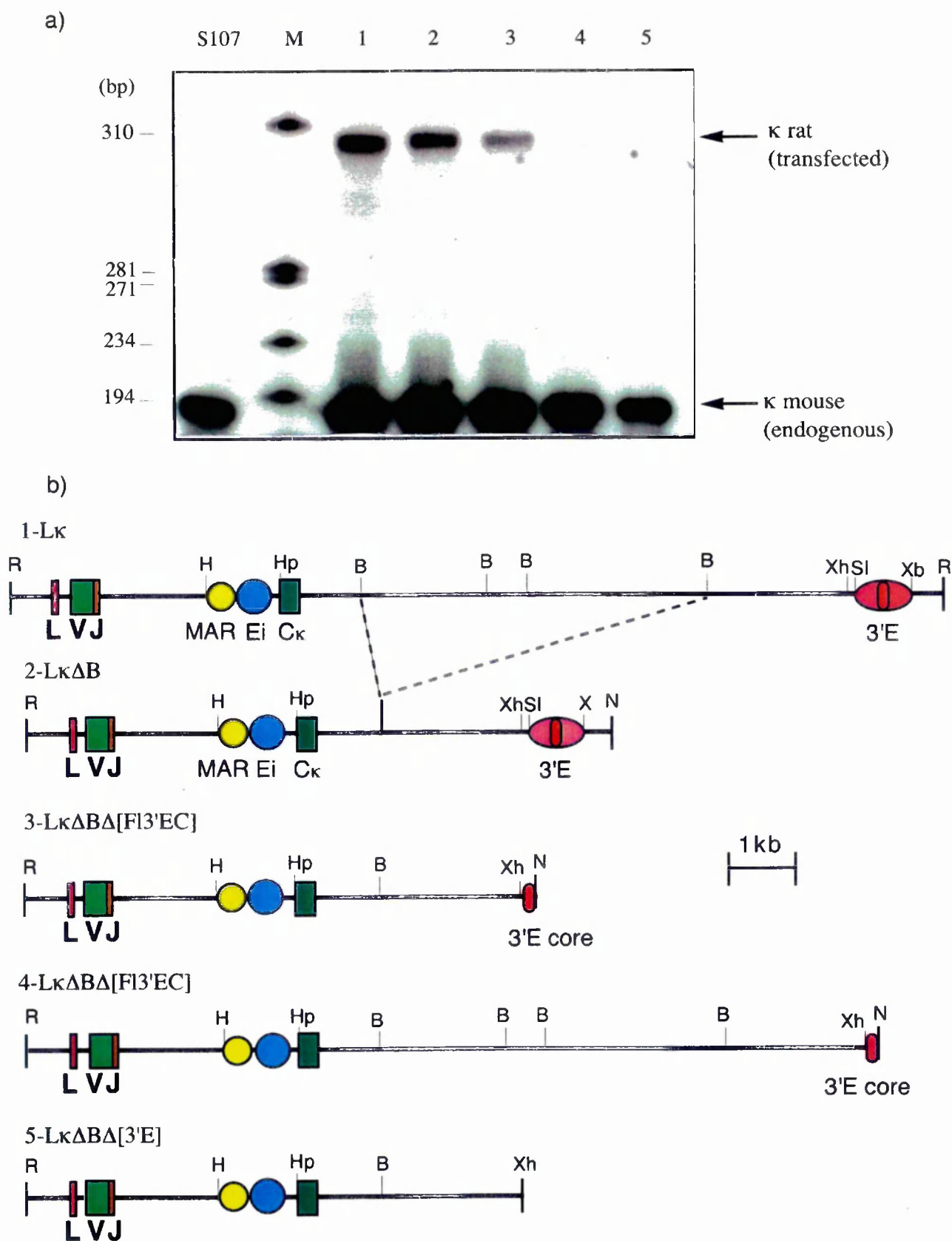


Figure 3.10 Expression of the κ transgenes transfected into S107 cells. a) Ribonuclease protection assay of transcription from the transfected κ gene of S107 cells stable transfected with the constructs shown in panel b. The experiment was performed several times, and representative data are presented. The position of size markers (a Hae III digest of ϕ X174) are shown in the lane marked "M". b) The map of the constructs used is shown.

3.5.4 Transgene κ chain expression in stably transfected S107 cells

The role of the whole 3' enhancer in the immunoglobulin gene expression was previously analysed in transgenic mice carrying a rearranged κ gene with or without the κ 3' enhancer. Its omission leads to 20-40 fold lower expression of the κ transgene and to poor allelic exclusion (Meyer et al., 1990). However, the experiments that lead to the definition of the 3' enhancer core were done in transient transfection assays using reporter genes. The contribution of the 3' enhancer core on its own to the κ expression therefore remains unresolved. To address this question the role of the core enhancer was tested in stably transfected S107 cells.

To study the contribution of the 3' enhancer core to the κ expression a series of stable transfection experiments was carried out using the S107 plasmacytoma cell line, which lacks functional NF- κ B. This cell line provides a good system to test the κ rat expression driven by the κ 3' enhancer and its derivatives without the interference of the κ intron enhancer activity. The presence of the SV40 enhancer in the vector is not able to compensate for the lack of the 3' enhancer or its flanking regions since the SV40 enhancer is very weak in these cells, possibly because of a dependency of SV40 enhancer activity in plasmacytoma cells on NF- κ B (Atchison and Perry, 1987).

A number of constructs with and without the 3' enhancer flanking regions were used for transfections. The intact L κ construct (Meyer et al., 1990) was used as a positive control. Two constructs (L κ Δ [F13'EC] and L κ Δ B Δ [F13'EC]) carrying the deletions of the 3' enhancer core flanking regions (see Materials and Methods for detailed description of the constructs) were prepared. A construct (L κ Δ B Δ [3'E]) without the 3'E was made as control. The short construct L κ Δ B (Yélamos et al., 1995), which carries the deletion of the DNA fragment between the BamH I sites 3' of the C κ , was also included in these experiments.

The plasmids were transfected into S107 cells and rat κ expression in pools of stable transfectants was measured by ribonuclease protection assays. Transcripts from the endogenous mouse κ gene of S107 partially hybridized with the rat κ probe and served as internal control (Fig. 3.10). The expression level was quantitated by densitometry and the values for κ rat were normalised for equal RNA amounts given by the endogenous mouse κ value. The ratio (rat κ / mouse κ) obtained for the L κ transfectant was considered as 100% of expression. It should be noted that the expression of the endogenous κ was higher than the expression of the transgenic κ rat in all the transfectants. This difference was also observed for a construct carrying the rearranged κ gene stably transfected into S194 cells (Blasquez et al., 1989). The relative rat κ expression for all the constructs tested is shown in Table 3.2.

Table 3.2 Expression of transfected and endogenous κ genes in S107.

Construct and controls	Titre* rat κ exp.	Titre \diamond mouse κ exp.	Relative rat κ exp. \downarrow (ELISA)	Relative rat κ exp. \downarrow (RNA prot. assay)
L κ	2.5	1:80	100%	100%
L κ Δ B	1.5	1:211	63%	80%
L κ Δ B[F13'EC]	6	1:111	30%	52%
L κ Δ [F13'EC]	7	1:143	20%	30%
L κ Δ B Δ [3'E]	No inhibition \dagger	1:81	<0.01%	0%
S107	No inhibition \dagger	1:500	<0.01%	0%
NQN1 B9.6	1:62	1:90		
P4C1 as κ chain	10 ⁻⁸ M	-		
Y3	0.5x10 ⁻⁹ M	-		
9E10 as κ chain	-	2.5x10 ⁻⁹ M		

* Titre of rat κ expression is given as the dilution (values <1) or concentration (values >1) of supernatant required to sequester 50% of the anti rat κ antibody used in the ELISA. For P4C1 and Y3 the values are given in molar concentration.

\dagger No inhibition was detected when supernatant was concentrated 10 times.

\diamond Titre of endogenous κ expression is given as the dilution of supernatant required to sequester 50% of the anti mouse κ antibody used in the ELISA. The value for 9E10 is given in molar concentration.

\downarrow Rat κ expression (rat κ / mouse κ) of L κ was defined as 100%, and all numbers are expressed relative to this value.

Supernatant from NQN1 B9.6, an anti-phOx hybridoma from L κ Δ [F1] (Betz, 1994), was included as a control in the ELISA.

P4C1 is the purified fraction I containing rat κ associated with IgG1 of P4C1 antibody (Suresh et al, 1986).

Y3 is a purified κ rat light chain (Bence-Jones protein).

9E10 is a purified mouse antibody containing a κ light chain.

The core enhancer was able to direct expression of the κ gene in stably transfected S107 cells, but only up to 30% of the activity of the full enhancer. In previous experiments the core enhancer was linked to the CAT reporter gene in transiently transfected S194 cells (Pongubala and Atchison, 1991), where it retained approximately 75% of the entire enhancer activity. It is possible that this discrepancy is due to the differences between the assays used to measure the expression.

The construct L κ Δ B Δ [F13'EC], in which the core enhancer has been brought closer to the C κ exon, gave a higher level of κ transgene expression (52%) than the L κ Δ [F13'EC] construct (30%). Differences were also found for the expression of L κ and L κ Δ B carrying the intact enhancer, but in this case the L κ Δ B construct expressed slightly less (80%) than the L κ one. As expected, expression was not detectable when the 3' enhancer was removed.

Rat κ expression in pools of stable transfectants was also measured by ELISA. The supernatant was pooled and the titres of rat and mouse κ were analysed (Table 3.2). The titres are given as the dilution or the concentration of the culture supernatant required to sequester 50% of the anti rat or mouse κ antibody used in the ELISA. The results are given as the mean of several experiments.

The core enhancer in the L κ Δ B Δ [F13'EC] construct retained only 30% of the expression while the core enhancer in the L κ Δ [F13'EC] construct only retained 20% of the expression. The short construct carrying the intact 3' enhancer (L κ Δ B) expressed slightly less (63%) than the L κ one. These data confirm the results obtained by the RNA protection assay.

To ensure that the expression levels seen by both ribonuclease protection assays and by ELISA using mixed cell populations are not biased due to differential outgrowth of specific clones, supernatant titres of rat and mouse κ in individual wells were analysed for specific clones by ELISA. The results reflect those seen in the ribonuclease protection assays and in ELISA with the pools (data not shown).

An intriguing feature of the inhibition of ELISA assays was detected in the controls. The concentration of rat κ forming part of an antibody (P4C1) required for 50% inhibition was 20 fold lower than that observed when the Bence-Jones free κ chain protein (Y3) was used. The reasons for this phenomenon are unknown but may affect the accuracy (but not the essential features) of the experiments described above.

3.6 Discussion

Seven transgenic lines were obtained to study the role of different DNA elements in hypermutation. In three of those lines a different approach was employed. In these lines the derived mice contain multiple, distinctly modified transgenes. This strategy allows

several transgenes to be tested at the same time, and one of them can serve as internal control for the others.

The configuration of the integrated transgenes in all lines was studied by Southern blotting. The analysis of the transgene copy number in the offspring revealed the intrinsic instability of inverted repeats structures present at so called "head-to-head" or "tail-to-tail" joining of transgene monomers. The instability was detected as deletions of those configurations (and DNA fragments flanking them). High rate of structural instability associated with the presence of an inverted repeat has been observed previously in two transgenic lines (Collick et al., 1995).

The analysis of the transgene expression in the mouse sera revealed that the same high levels of expression produced in Lk6 line was only achieved in LkNG line. The Lk(WT&7M) line, which carries only one intact copy of the Lk construct, the transgenic κ expression was about 2.5-10 fold lower than in the Lk6 and LkNG lines. This difference may be due to the lower copy number of the transgene in the Lk(WT&7M) line (5 copies in the first two lines).

In the Lk Δ [F13'EC] line, which lacks the 3' enhancer core flanking region, the transgenic κ expression was reduced but not as much as when the whole enhancer was deleted (Meyer et al., 1990; Betz et al., 1994). Even lower transgenic κ expression was observed when the transgene lacks either the 3' enhancer core or the intron enhancer.

The data showed here are in agreement with observations from transients and stable transfection assays using artificial reporter gene constructs as well as rearranged genomic transgenes (Blazquez et al., 1992; Fulton and Van Ness, 1993; Fulton and Van Ness, 1994). These studies proved the requirement of both enhancers for maximal gene expression in a synergistic manner.

The transgene expression in Lk Δ [MAR] line was also reduced. This is in line with previous observations in stably transfected plasmacytoma cells harbouring a construct containing the κ locus carrying deletion of the MAR (Blasquez et al., 1989). In this study it was also shown that the effect is eliminated when the same construct was coupled to a transcribed selectable marker. In addition, transient expression assays employing vectors containing linked active genes failed to detect any effect of deleting the MAR on RNA levels (Bergman et al, 1984; Queen and Stafford, 1984). These observations suggest that the role of the MAR may be to set up a favourable chromatin environment for transcription, perhaps by targeting the chromatin to a specific subnuclear location enriched in the transcription machinery.

The results for the expression of the Lk Δ [MAR] and Lk Δ [Ei] transgenes in the serum of the mice are supported by the expression analysis in NS0 cells which have been stably transfected. This is not in agreement with previous results for the transgenic line Lk Δ [Ei/MAR] (Betz et al., 1994) where both Ei and MAR were deleted and the expression of the immunoglobulin gene was not affected. This discrepancy could be

caused by a site-integration and gene dosage effects. Thus, L κ Δ [Ei/MAR] transgenes maybe have been integrated in a actively open chromatin locus, which compensates the deletion of MAR. On the other hand a high copy number and/or rearrangement within the L κ Δ [Ei/MAR] transgenes may lead to higher levels of transgene expression.

The experiments performed with the S107 cells, which have been stable transfected, showed that the core of the 3' enhancer was able to drive κ expression in the absence of a functional NF- κ B. The 3' enhancer core retained up to 30% of the whole enhancer activity as measured by RNA protection assays and up to 20% measured by the protein present in the supernatants of the transfected cells. This value increased to 52% and 30% respectively when the 5kb fragment between the BamHI sites was removed.

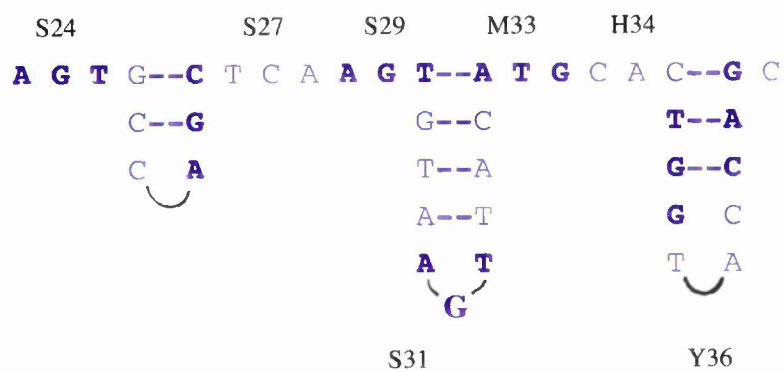
This increase however was not observed when the intact 3' enhancer was tested. On the contrary, the expression level obtained for the L κ Δ B construct was lower than that observed for the L κ construct in both assays. One possible reason for this observation could be the presence of a silencer in the deleted fragment which acts in conjunction with the 3' enhancer core but not with the full enhancer.

Transcriptional suppression of Ig expression was also observed when the IgH 3' enhancer was brought closer to the μ intron enhancer. This effect was also observed when the 3' enhancer was placed 5' end of the V gene, but not when the intron enhancer was deleted in the 3'E at 5' end of the V construct. These results indicate that the transcriptional suppression can not be ascribed to the proximity of the transcription initiation site, but rather to the close juxtaposition of the μ intron enhancer (Mocikat et al., 1995). Therefore, the IgH 3' enhancer may be a novel type of *cis*-regulating element. It can cooperate with the Ig μ intron enhancer, but its synergy is differentially regulated by the distance separating the two elements. These effects are similar to the ones found here for the κ 3' enhancer and the 3' enhancer flanking sequences may therefore be necessary for cooperation.

A distance effect in the transgene κ expression have been observed also by Meyer et al. (1990). The same L κ construct used here and another version L κ Δ B construct (carrying deletions of the sequences between the EcoRV sites) containing the intact enhancer were tested for κ expression in transfected NS0 cells. In these experiments the construct with the 3' enhancer closer to the V elements gave higher level of κ transgene expression than the L κ construct. However κ transgene expression in NS0 cells may be affected by the SV40 promoter, which is also able to direct the transgene expression. Therefore this may not provide a good example to compare with S107 transfectants.

.....

S24 A25 S26 S27 S29 V30 S31 Y32 M33 H34 W35 Y36 Q37
AGT GCC **AGC** TCA **AGT** GTA AGT TAC **ATG** CAC **TGG** TAC **CAG**



Chapter Four Influence of DNA structure and primary sequence in somatic hypermutation

4.1 Introduction

The CDR1 has a large accumulation of hotspots and is usually more mutated than other regions. The hotspots in the CDR1 of the V κ Ox1 gene (the most frequently used in the immune response against the hapten 2-phenyl-axazol-5-one) has been studied extensively. It is well known which hotspots are antigen selected or which are intrinsic to the hypermutation process. The V κ Ox1 is therefore a very attractive system to study the intrinsic features of the mutational machinery.

The intrinsic hotspots in CDR1 of the V κ Ox1 gene (second base of Ser31 and third base of Ser26) exemplify the consensus: PuGPyT/A, with G being most frequently mutated. However, not all the hotspots conform to this consensus. For example, Val30 (third base) and Tyr36 (second base) are intrinsic hotspots that have in common a 'TA' core with A being the mutated base.

Intrinsic hotspots are only in part created by local DNA sequence since they are not found in all the positions conforming to the hotspot consensus. Thus, other structural features, such as palindromes, may well be important.

The analysis of the CDR1 in the V κ Ox1 gene reveals several possible palindromic structures which may give rise to hairpin loops. The thermodynamically most stable are shown in Fig. 4.1. These different structures may be in a equilibrium. The palindromic structure which contains the Ser31 is very stable, and the hypermutable G is at the tip of the postulated hairpin.

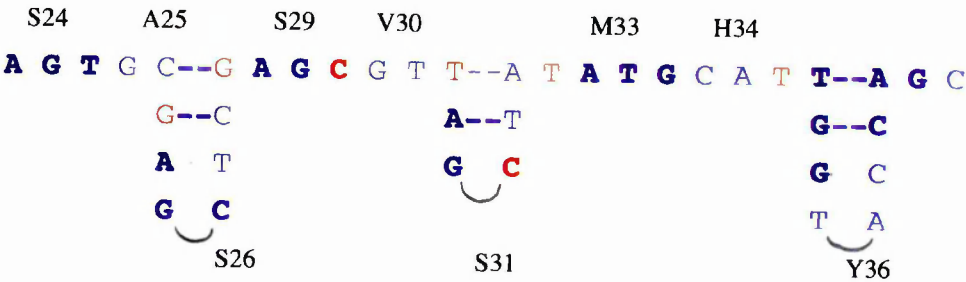
Figure 4.1 DNA sequence and predicted secondary structure in the CDR1 of the V κ Ox1 gene (opposite page). The most stable hairpin loops are shown. Triplets are alternatively overprinted for a easier amino acid identification.

a)

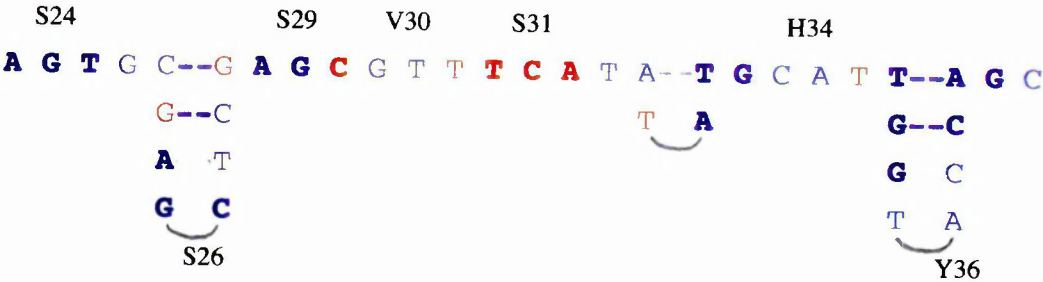
CDR1

	S24	A25	S26	S27	S29	V30	S31	Y32	M33	H34	W35	Y36	Q37
WT	AGT	GCC	AGC	TCA	AGT	GTA	AGT	TAC	ATG	CAC	TGG	TAC	CAG
7M	AGT	GCG	AGC	TCG	AGC	GTT	AGC	TAT	ATG	CAT	TGG	TAC	CAG
7M/TCA	AGT	GCG	AGC	TCG	AGC	GTT	TCA	TAT	ATG	CAT	TGG	TAC	CAG

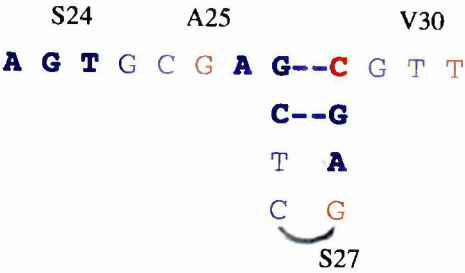
b) 7M



c) 7M/TCA



d)



4.1.1 Modifications introduced in the CDR1 of the V κ transgenes

To assess the effect of the local DNA secondary structure and the primary sequence in the position of hotspots several silent mutations were introduced in the CDR1 of the V κ Ox gene of the κ transgene (see Materials and Methods). These mutations were designed to change the DNA sequence in the hotspot consensus and to disrupt or modify the secondary structure in the CDR1. The DNA sequence and the proposed secondary structure of two mutants are shown in Fig. 4.2.

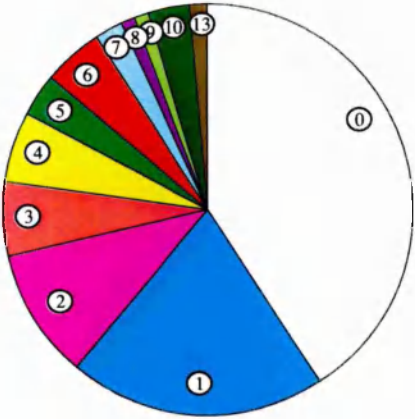
The contribution of those mutations on hotspots was studied in three transgenic mouse lines: L κ NG (which contains L κ Δ B and L κ Δ B7M transgenes among others), L κ (WT&7M) (which contains L κ and L κ 7M transgenes) and L κ 7M/TCA& Δ [Li] (which contains L κ Δ B7M/TCA and L κ Δ [Li] transgenes).

Hypermutation was studied in germinal centre PNA^{hi} B cells from Peyer's patches (collected from several mice about 5 month old), and in hybridomas generated from phOx-CSA immunized mice (see Materials and Methods).

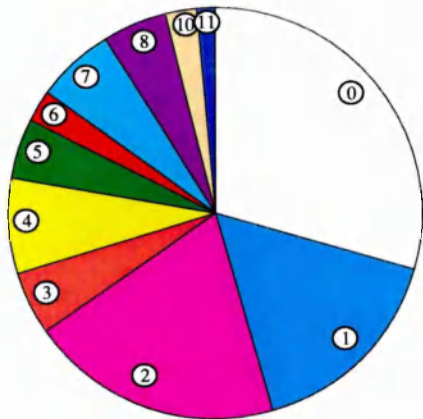
Hypermutation was analysed as described in Materials and Methods by amplification with the oligonucleotides VKOXFOR and VKOXBACK (Appendix II) using Taq polymerase. The sequencing primer was the oligonucleotide JK5. The PCR error frequency (0.4×10^{-3} per base) was determined by sequencing L κ Δ B and L κ Δ B7M transgenes from hybridomas of the L κ NG line.

In the L κ 7M/TCA& Δ [Li] line PCR amplification of the L κ Δ B7M/TCA transgene was carried out with the oligonucleotides VKOXFOR and BEALEADER. These primers allow a specific amplification of the 7M/TCA copies as BEALEADER hybridises in the leader region that was deleted in the Δ [Li] construct (see Appendix II). The amplification was done with Pfu polymerase, and the PCR error frequency (0.13×10^{-3} per base) was determined by sequencing L κ 7M/TCA transgenes from hybridomas.

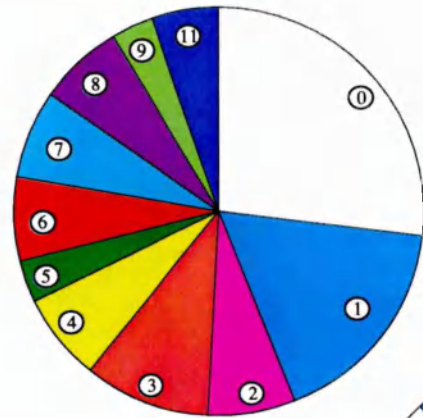
Figure 4.2 The predicted secondary structure in the CDR1 of the V κ Ox1 mutated gene (opposite page). The possible most stable hairpin loops are shown. Triplets are alternatively overprinted for a easier amino acid identification. a) CDR1 DNA sequence of the wild type and the mutants. The mutations are shown in red. The underlined sequences correspond to the hotspot consensus (PuGPyT/A). b) Secondary structure of the mutant 7M. c) Secondary structure of the mutant 7M/TCA. d) An alternative hairpin of both mutants.



LkWT in Lk(WT&7M)



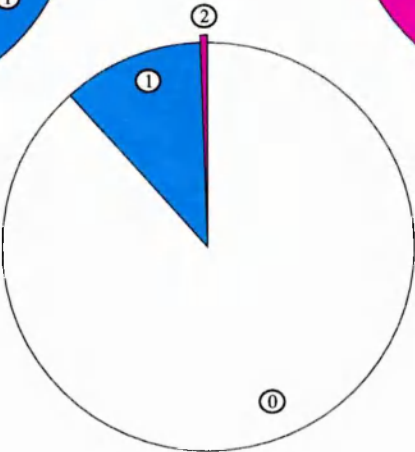
Lk7M in Lk(WT&7M)



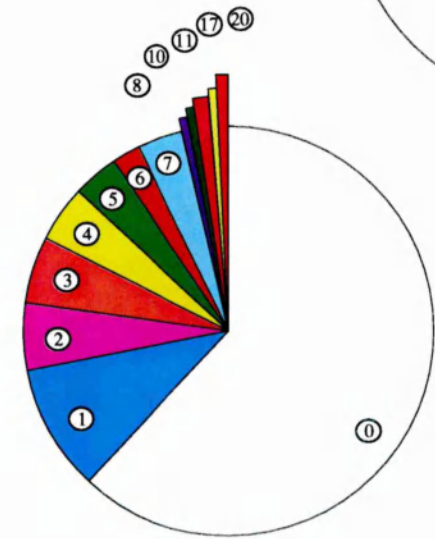
LkΔBWT in LkNG



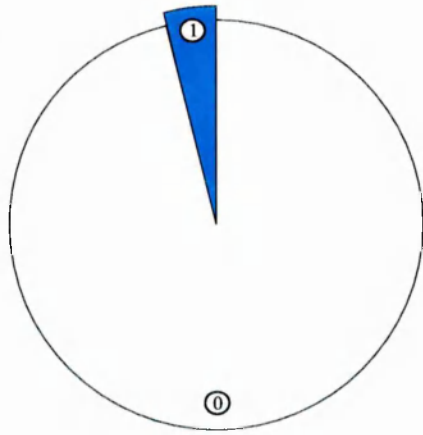
LkΔB7M in LkNG



PCR error (Taq polymerase)



LkΔB7M/TCA



PCR error (Pfu polymerase)

4.2 The κ 3' enhancer drives hypermutation from a position 5' of the target area

L κ (WT&7M) carries two V κ segments (WT and 7M) of which only one belongs to an intact transgene copy. The other is part of a truncated copy that contains the Ei/MAR-C κ region but not the 3' enhancer. Southern blot analysis of tail's DNA digested with several restriction enzymes was performed to try to identify the complete copy. However, due to the symmetry of the transgene this was not possible.

The sequence analysis of the transgenes in the L κ (WT&7M) line revealed that both transgenes had been substrate for hypermutation (Table 4.1 and Fig. 4.3). The sequences with one or more mutations of the L κ and L κ 7M are shown in the Appendices III E and F respectively.

The number of mutations per clone found in 7M and WT copies is very similar (3.5 and 3.7 respectively). Amongst the 7M clones sequenced 29.6% were unmutated. A higher number of WT clones (40.9%) were unmutated. The overall frequency of mutations in the 7M copy is slightly higher (9.2×10^{-3} mutations per base pair) than that found in the WT copy (7.4×10^{-3} mutations per base pair).

However, this is not the case in the L κ NG transgenic line which also carries 7M and WT transgenes. On the contrary, in the L κ NG line the 7M copies mutate at a frequency slightly lower (7.1×10^{-3} mutations per base pair) than the WT copies (11.8×10^{-3} mutations per base pair) and the number of mutations per clone in the WT is higher than in the 7M copies (4.6 and 2.9 respectively).

Figure 4.3 Mutations in the transgenic V regions (opposite page). Each pie chart depicts the proportion of sequences with the specified number of mutations. A pie chart showing the control for PCR errors is included for comparison. PCR with Taq polymerase (L κ , L κ 7M, L κ Δ B, L κ Δ B7M), the PCR error was measured by sequencing L κ Δ B and L κ Δ B7M transgenes from characterized hybridomas. PCR with Pfu polymerase (L κ 7M/TCA), the PCR error was measured by sequencing L κ Δ B7M/TCA transgenes from characterized hybridomas.

Table 4.1 Lκ, Lκ7M, LκΔB, LκΔB7M and LκΔB7M/TCA transgene mutations in PCR clones derived from Peyer's patch B cells (PNA^{hi}).

Transgene	Clones [#]			Mutations [*]		Mut./clone [□]		Mutations/10 ³ base pair [¶]	
	Total	=1	>1	Total		≥1	>1	MFt	MF(>1)
Lκ(WT)	88	18	34	184		3.5	4.9	7.4 (7)	17.3
LκΔB(WT)	59	10	33	197		4.6	5.7	11.8 (11.4)	20.1
WT (Average)	147	28	67	381		4	5.2	9.2 (8.8)	18.7
Lκ(7M)	81	13	44	211		3.7	4.5	9.2 (8.8)	16
LκΔB(7M)	26	6	12	52		2.9	3.8	7.1 (6.7)	13.6
7M (Average)	107	19	56	263		3.5	4.4	8.7 (8.3)	15.4
PCR error ^Ø	245	27	1	29		0.9	2	0.4	7
LκΔB7M/TCA	172	17	48	267		4.1	5.2	5.5 (5.4)	18.5
PCR error [†]	52	2	0	2		1	0	0.13	0
Lκ3 [‡]	76	9	44	267				12	21
Lκ6(5 month) [‡] 37				123				11	
PCR error								0.7	

[#] Total number of PCR-generated clones and the number of those carrying one (=1) or more than one (>1) mutations.

^{*} Sum of mutations in all clones.

[□] The number of mutations per clone was calculated using either the number of clones carrying mutations (≥1) or only those carrying more than one mutation (>1).

[¶] The frequency of mutations per 10³ base pair was calculated using either the total number of clones (MFt) or only those clones carrying more than one mutation (MF(>1)). The values after correction of the PCR error are given in parentheses.

[†] Clones obtained from hybridomas of Lκ7M/TCA&Δ[Li] transgenic line.

^Ø Clones obtained from hybridomas of LκNG transgenic line.

(Average) Data from either WT or 7M transgenes which originate from LκNG and Lκ(WT&7M) lines were used for the average.

[‡] Data from González-Fernández and Milstein, 1993.

[‡] Data from González-Fernández et al., 1994a.

The 3' enhancer of the neighbour copy must be able to act from the 5' position with respect to the V κ region because both V κ segments mutate. It was previously shown that the 3' enhancer can drive hypermutation both when moved closer to the V-C elements and when present in the opposite orientation (Betz et al., 1994). The results shown here indicate that the downstream position of the 3' enhancer with respect to the promoter elements is not important for hypermutation.

4.3 The modifications in the DNA sequence and secondary structure did not affect the rate or the nucleotide substitution preferences of hypermutation

The hypermutation analysis of the modified 7M and 7M/TCA transgenes showed that they mutate at a level comparable to that observed in the L κ (WT) transgenes (Table 4.1, Fig. 4.3). The sequences of the L κ Δ B, L κ Δ B7M and L κ Δ B7M/TCA with one or more mutations are shown in the Appendices III G, H and I respectively.

As described in the previous section, differences in the overall mutation frequency and mutation per clone between the WT and 7M transgenes in the L κ (WT&7M) and L κ NG lines were found. This is also observed in the hybridomas generated from both lines (section 4.6). The variations are not caused by the mutations introduced in the CDR1 of the 7M V κ gene. They rather may be produced by a differential targeting of the transgenes.

The reasons for this differential targeting are unknown. It may be a result of a poorer transcription of some transgene copies. In the L κ (WT&7M) line this could be caused by the different position of the 3' enhancer in one of the copies. In the L κ NG line may contain truncated copies, as the number of transgene copies estimated from the Southern blot probed with the J-C intron fragment did not correspond to the number estimated when the 3'E probe was used (see Chapter 3).

The mutation frequency amongst L κ 7M clones carrying more than one mutation (16×10^{-3} mutations per base pair) is very similar to that observed in L κ WT (17.3×10^{-3} mutations per base pair) clones in the L κ (WT&7M) line (Table 4.1).

In the L κ NG line the frequency of mutations of the WT and 7M transgenes are also comparable, although the mutation frequency of clones carrying more than one mutation in the wild type (20.1×10^{-3} mutations per base pair) is slightly higher than in the mutated transgene (13.6×10^{-3} mutations per base pair).

In the L κ 7M/TCA& Δ [Li] line the clones carrying more than one mutation in L κ 7M/TCA transgene mutate with a frequency of 18.5×10^{-3} mutations per base pair.

However, when the overall mutation frequency is taken into consideration, only the L κ Δ B(WT) has a similar frequency (11.8×10^{-3}) to the one found previously in the L κ 3 and L κ 6 transgenic lines ($11-12 \times 10^{-3}$ mutations per base pair). The largest difference was found for L κ 7M/TCA transgenes, where the mutation frequency (5.5×10^{-3} mutations per base) was 2 times lower. This is due to the low percentage of mutated

clones (Fig 4.3). About 59 to 73% of the sequences were mutated in all the transgenes analysed except in Lκ7MTCA where only 38% of the sequences carried mutations.

The proportion of mutated clones carrying more than one mutation is very similar for all the lines (65-77%). Therefore, the lower overall mutation frequency observed in Lκ7M/TCA transgenes can not be due to a decreased number of highly mutated clones.

The nucleotide substitution preferences of mutations in Lκ, Lκ7M and Lκ7M/TCA Peyer's patches (Table 4.2) is comparable to the previously defined substitution data from Lκ6 Peyer's patches (González-Fernández and Milstein, 1993). The predominant substitutions show the typical bias of transitions over transversions (Table 4.2) as well as the typical bias against mutations in pyrimidines residues on the coding strand.

Thus there is no feature specific neither to the secondary structure nor the DNA sequence in the CDR1 that is essential for recruiting hypermutation.

Table 4.2 Nucleotide substitution in Lκ and Lκ mutated transgenes from Peyer's patch PNA^{hi} B cells.

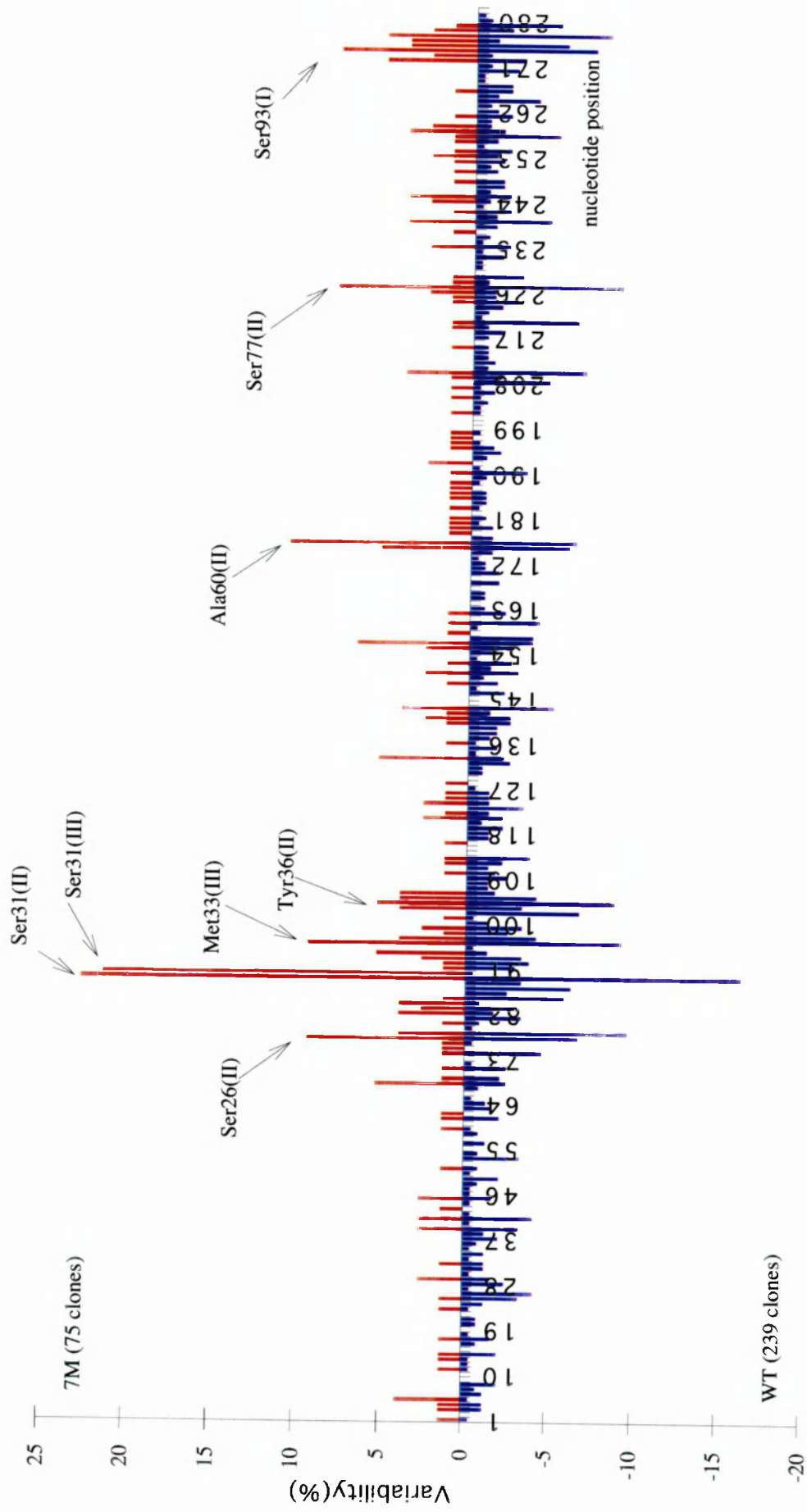
Lκ		To		G	Total mutations*	Corrected % ¶
From	T	C	A			
T	-	30 (0.58)	15 (0.29)	7 (0.13)	52	0.15
C	80 (0.81)	-	8 (0.08)	10 (0.10)	98	0.23
A	34 (0.28)	25 (0.20)	-	64 (0.52)	123	0.32
G	8 (0.07)	14 (0.13)	86 (0.80)	-	108	0.30
.....						
Lκ7M		To		G	Total mutations*	Corrected % ¶
From	T	C	A			
T	-	18 (0.40)	12 (0.27)	15 (0.33)	42	0.18
C	61 (0.83)	-	4 (0.05)	8 (0.11)	73	0.25
A	21 (0.29)	12 (0.16)	-	40 (0.55)	73	0.28
G	5 (0.07)	15 (0.21)	52 (0.72)	-	72	0.28
.....						
Lκ7M/TCA		To		G	Total mutations*	Corrected % ¶
From	T	C	A			
T	-	13 (0.31)	18 (0.43)	11 (0.26)	42	0.16
C	45 (0.78)	-	2 (0.03)	10 (0.17)	57	0.19
A	34 (0.34)	22 (0.22)	-	45 (0.44)	101	0.38
G	9 (0.13)	15 (0.22)	43 (0.64)	-	67	0.26

The figures in the three sections give, for each of the four bases in the Vκ sequence, the number of times and the proportional distribution (in parenthesis) with which it is mutated to one of the other three bases.

Lκ data originate from Lκ and LκΔB transgenes. Lκ7M originate from Lκ7M and LκΔB7M transgenes.

* Total number of times that the four bases were found mutated.

¶ The proportion of mutations in each base corrected for the base composition of the region analysed.

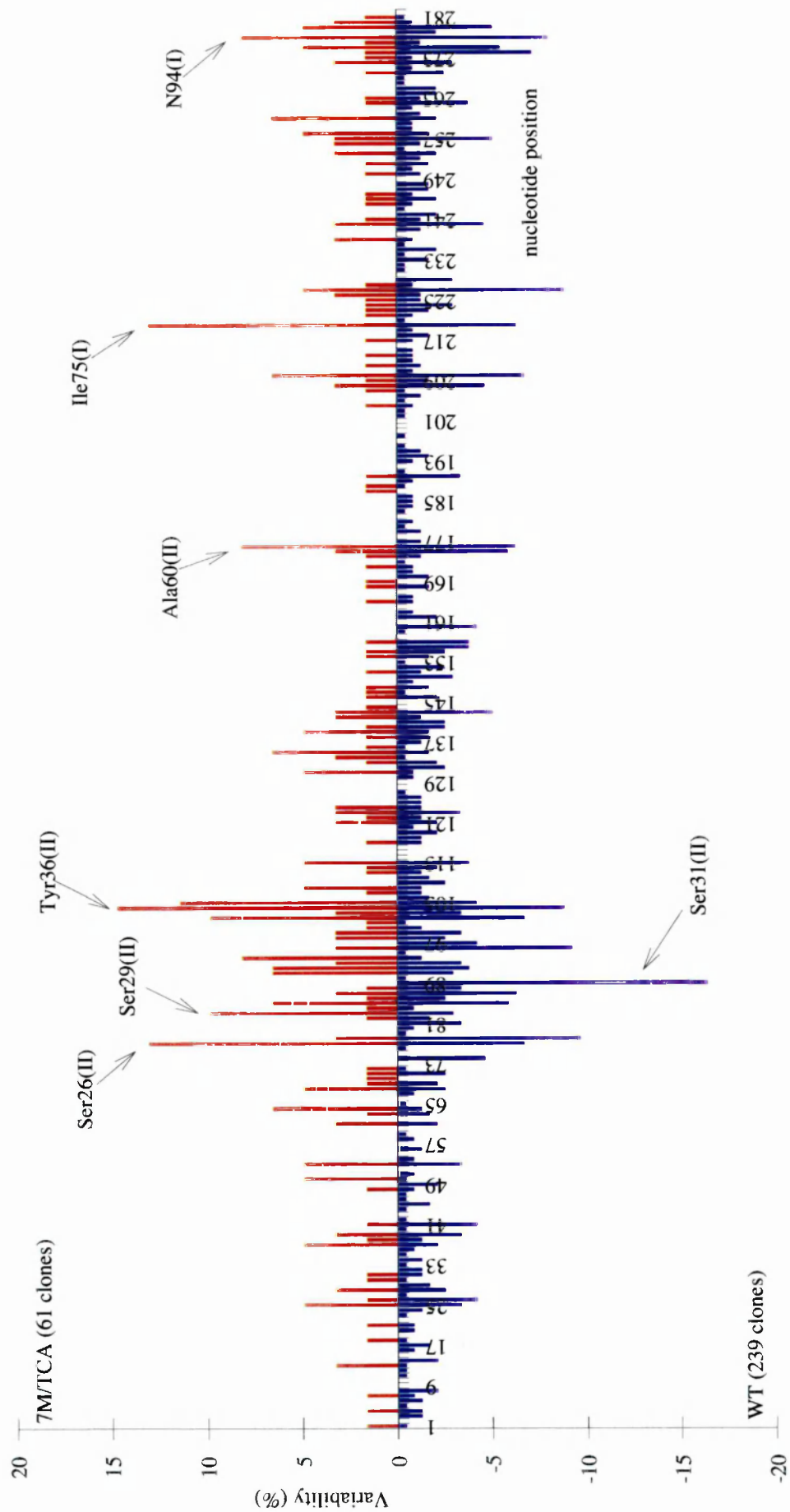


4.4 The DNA sequence has a major influence on hotspots

The distribution of somatic mutations along the modified and wild type L κ transgenes is shown in Fig. 4.4 and 4.5, where the variability is given as the percentage of sequences that carry a mutation at each particular nucleotide of the V κ region. The segment analysed includes 282bp using as the starting point the first coding nucleotide for the mature variable region.

A large compilation of transgenic L κ (WT) sequences from Peyer's patch B cells was constructed with 95 sequences from L κ and L κ Δ B (see Appendices III E and G); 107 sequences from L κ 6 and L κ 3 (González-Fernández and Milstein, 1993; González-Fernández et al., 1994a); 20 sequences from L κ in ENG (González-Fernández, personal communication) and 17 sequences from L κ Δ [3'Fl] (Betz et al., 1994). This extensive database of L κ (239 sequences with a total of 1112 mutations) was used for comparison to assess whether the distribution of mutations is affected when the L κ transgene carries a modification.

Figure 4.4 Distribution of mutations on L κ 7M and L κ Δ B7M transgenes (opposite page). The distribution of mutations amongst the transgenic V regions clones isolated from Peyer's patch PNA^{hi} B cells from L κ NG and L κ (WT&7M) transgenic lines is shown. Variability at each nucleotide position is computed as the percentage of sequences carrying one or more mutations which are mutated at that position. Nucleotides are numbered from the first base of the initiator ATG codon. Hotspots are indicated with the name and position of the amino acid. The distribution of mutations on L κ transgenes is provided for comparison.



The compilation of mutations in the Lκ7M and Lκ7M/TCA transgenes shows the expected non random distribution of mutations and the presence of hotspots preferentially around CDR1. It is possible to find some of the most dominant intrinsic hotspots previously identified in the large data base for the same gene: the second base of Ala60, the second base of Ser77, the first base of Ser93, the first base of N94. However, in the modified CDR1 substantial differences were found.

4.4.1 A new intrinsic hotspot in Ser31

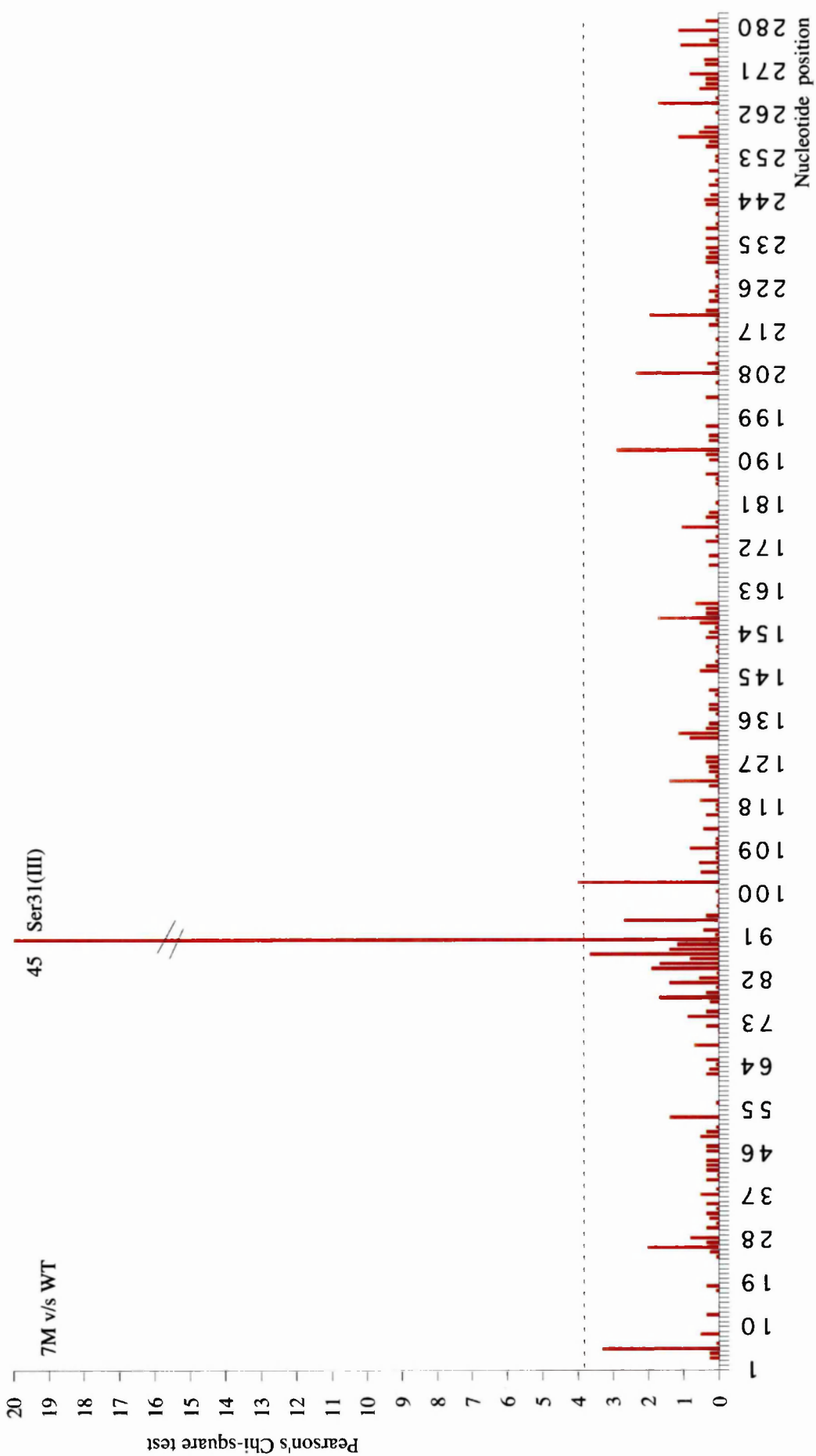
The percentage of sequences that carry a mutation at each particular nucleotide in the Vκ region of the 7M mutant was compared with that in the wild type. This analysis disclosed several differences in CDR1 region associated with the DNA sequence modifications. Higher number of mutations in Ser31(II) were observed. The second base of Ser31 AGT (the major recognised hotspot in several V genes) was found mutated in 16.3% of the clones sequenced in the Vκ wild type database, while the same position in the Vκ 7M, Ser31 AGC was found mutated in 22.6% of the clones (Fig. 4.4 and Table 4.3). A much larger change was found in the third base of Ser31. Contrasting with the 0.4% found in wild type AGT, 21.3% of the clones were mutated in the third base of the mutant Ser31 AGC. Differences were also observed in Met33(I) and Val30(III) and W35(III).

Figure 4.5 Distribution of mutations on LκΔB7M/TCA transgene (opposite page). The distribution of mutations amongst the transgenic V regions clones isolated from Peyer's patch PNA^{hi} B cells of the Lκ7M/TCA&Δ[Li] transgenic line is shown. Variability at each nucleotide position is computed as the percentage of sequences carrying one or more mutations which are mutated at that position. Nucleotides are numbered from the first base of the initiator ATG codon. Hotspots are indicated with the name and position of the amino acid. The distribution of mutations on Lκ transgenes is provided for comparison.

Table 4.3 Mutations, variability and Chi-square test for the CDR1 nucleotides in the Lκ WT and the mutated Lκ transgenes.

	WT	mut.	V(%)	7M	mut.	V(%)	χ^2	TCA	mut.	V(%)	χ^2
A25	G	0	0		0	0	–		0	0	–
	C	11	4.6		1	1.3	0.9		0	0	1.7
	C	0	0	G	1	1.3	0.4	G	0	0	–
S26	A	1	0.4		1	1.3	0		0	0	0.5
	G	16	6.6		7	9.3	0.3		8	13.1	1.9
	C	23	9.6		3	4	1.7		2	3.2	1.8
S27	T	1	0.4		0	0	0.4		0	0	0.5
	C	2	0.8		1	1.3	0.1		0	0	0
	A	8	3.3	G	0	0	1.4	G	0	0	1.0
S29	A	4	1.6		3	4	0.5		1	1.6	0.3
	G	7	2.9		2	2.6	0.1		6	9.8	4.0
	T	2	0.8	C	3	4	1.9	C	1	1.6	0.0
V30	G	14	5.8		1	1.3	1.6		4	6.5	0.0
	T	6	2.5		0	0	0.8		1	1.6	0
	A	15	6.2	T	0	0	3.6	T	2	3.2	0.3
S31	A	8	3.3		0	0	1.4	T	1	1.6	0.1
	G	39	16.3		17	22.6	1.1	C	0	0	10
	T	1	0.4	C	16	21.3	44.8	A	0	0	0.5
Y32	T	7	2.9		1	1.3	0.1		4	6.5	0.9
	A	9	3.7		1	1.3	0.4		4	6.5	0.3
	C	8	3.3	T	2	2.6	0	T	2	3.2	0.1
M33	A	3	1.2		4	5.3	2.7		5	8.1	6.5
	T	1	0.4		0	0	0.4		0	0	0.5
	G	22	9.2		7	9.3	0		2	3.2	1.5
H34	C	10	4.1		3	4	0.7		0	0	1.5
	A	6	2.5		1	1.3	0		2	3.2	0
	C	8	3.3	T	2	2.6	0	T	2	3.2	0.1
W35	T	3	1.2		0	0	0.1		1	1.6	1.1
	G	1	0.4		1	1.3	0		1	1.6	0
	G	16	6.6		0	0	3.9		6	9.8	0.3
Y36	T	8	3.3		3	4	0		2	3.2	0.1
	A	21	8.7		4	5.3	0.5		9	14.7	1.3
	C	10	4.1		3	4	0.1		7	11.4	3.5
A60	G	14	5.8		4	5.3	0		2	3.2	0.2
	C	15	6.2		8	10.6	1.0		5	8.1	0
	T	3	1.2		0	0	0.1		0	0	0.0
S77	A	3	1.2		2	2.6	0.1		2	3.2	0.3
	G	21	8.7		6	8	0		3	4.9	0.5
	C	2	0.8		1	1.3	0.1		1	1.6	0.0

The number of mutations (mut.), variability (V), and the results for the Pearson's Chi-square test (χ^2) for each nucleotide in the CDR1 region of the LκWT, Lκ7M and Lκ7M/TCA are shown. Other major hotspots are also included for comparison. The wild type amino acid and DNA sequence is indicated (WT). The silent mutations carried by the mutants 7M (7M) and 7M/TCA (TCA) are shown overprinted. The variability was calculated as the percentage of sequences that carry a mutation at each particular nucleotide. The Pearson's Chi-square test (on 1 degree of freedom) was calculate for the number of mutations of each mutant versus wild type. For significance at the 5% level, the Chi-square has to be greater than 3.84.



In order to assess whether these differences were significant a Pearson's Chi-square test (on 1 degree of freedom) was performed with the 7M database versus the WT database. The results show that the differences found are significant only for the third base of Ser31 AGC (Fig. 4.6, Table 4.3).

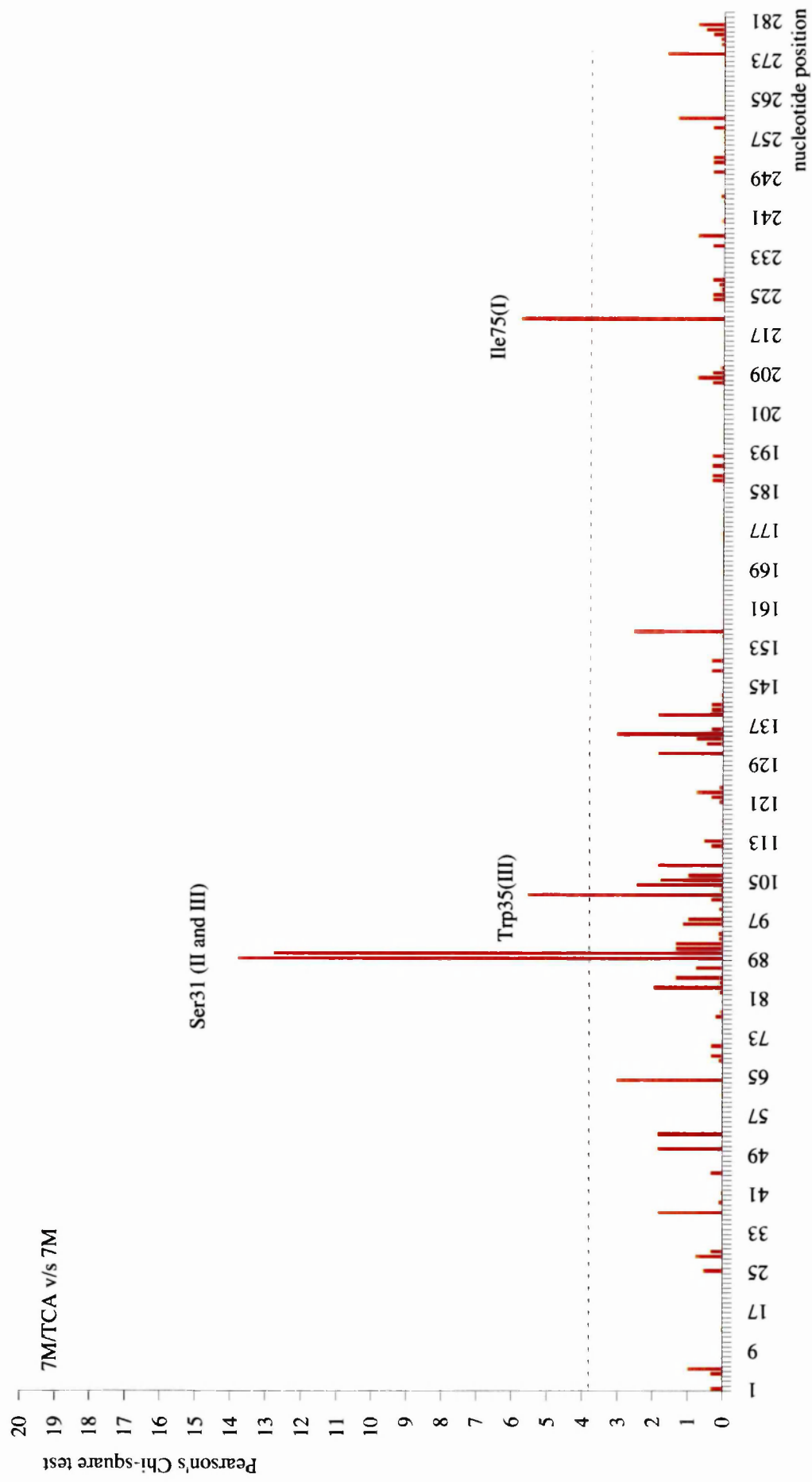
Therefore, Ser31 as AGC has not altered the major hotspot (second base). However the mutation produced an even more "attractive" motif (AGCT), which creates a new hotspot (in the third base). The maintenance of the G as hotspot and formation of a new intrinsic hotspot in Ser31 as AGC occurs in spite of the fact that the relevant hairpin is more unstable than that predicted for the wild type. The major hotspots in other regions of the V gene, like Ala 60 and Ser77 were not significantly altered.

4.4.2 Ser 31 as TCA is not a hotspot

When the AGT triplet in Ser31 was changed to TCA the classic major hotspot of the V genes did no longer exist. No mutations were found in the second or third bases and only one mutation was found in the first base of Ser 31 as TCA in 61 clones sequenced (Fig. 4.5 and Table 4.3). This contrasts with the 16.3% of clones carrying the Ser31 (second base) mutated in the WT V κ gene. Along with this change other differences were also observed in Ser29(II) and Met33(I) and no large differences were found in the other hotspots outside the CDR1. The significance of these difference was assessed with the Pearson's Chi-square test and the results are shown in Fig. 4.7.

Figure 4.6. Pearson's Chi-square statistic test for the differences (increase/decrease) in the mutability resulting from local base substitutions in 7M v/s WT.

The test was performed for each position along the V κ gene in the wild type and the 7M mutant for the clones sequenced (239 and 75 respectively) with the number of mutations found. For significance at the 5% level, the Chi-square need to be greater than 3.84 (shown by a dashed line).



The absence of mutations in Ser31 as TCA was significant confirming that the sequence motif has an important role in the origin of the hotspots. This finding is in line with previous observations that Ser as TCN triplets (N:A/T/C/G) mutated less than the AGY (Y:T/C) encoded ones (Schwager et al., 1989; Reynaud et al., 1995; Wagner et al., 1995). Those studies showed a strong bias for Ser as AGY to be used in the CDR regions and Ser as TCN to be used in the frameworks. The results shown here indeed confirmed that the ground for this bias are the motif preferences of the mutational machinery. They support the idea that during the germline V genes evolution they have evolved in response to selection for sequences which favour mutability in the CDRs and less favourable ones in the frameworks.

The comparison of 7M and 7M/TCA databases using the Pearson's Chi-square test (Fig. 4.8) revealed that basically the distribution of mutations in both mutants are similar except in the second and the third base of Ser31 due to the lack of mutations in Ser31(II) as TCA and the new hotspot in Ser31(III) as AGC.

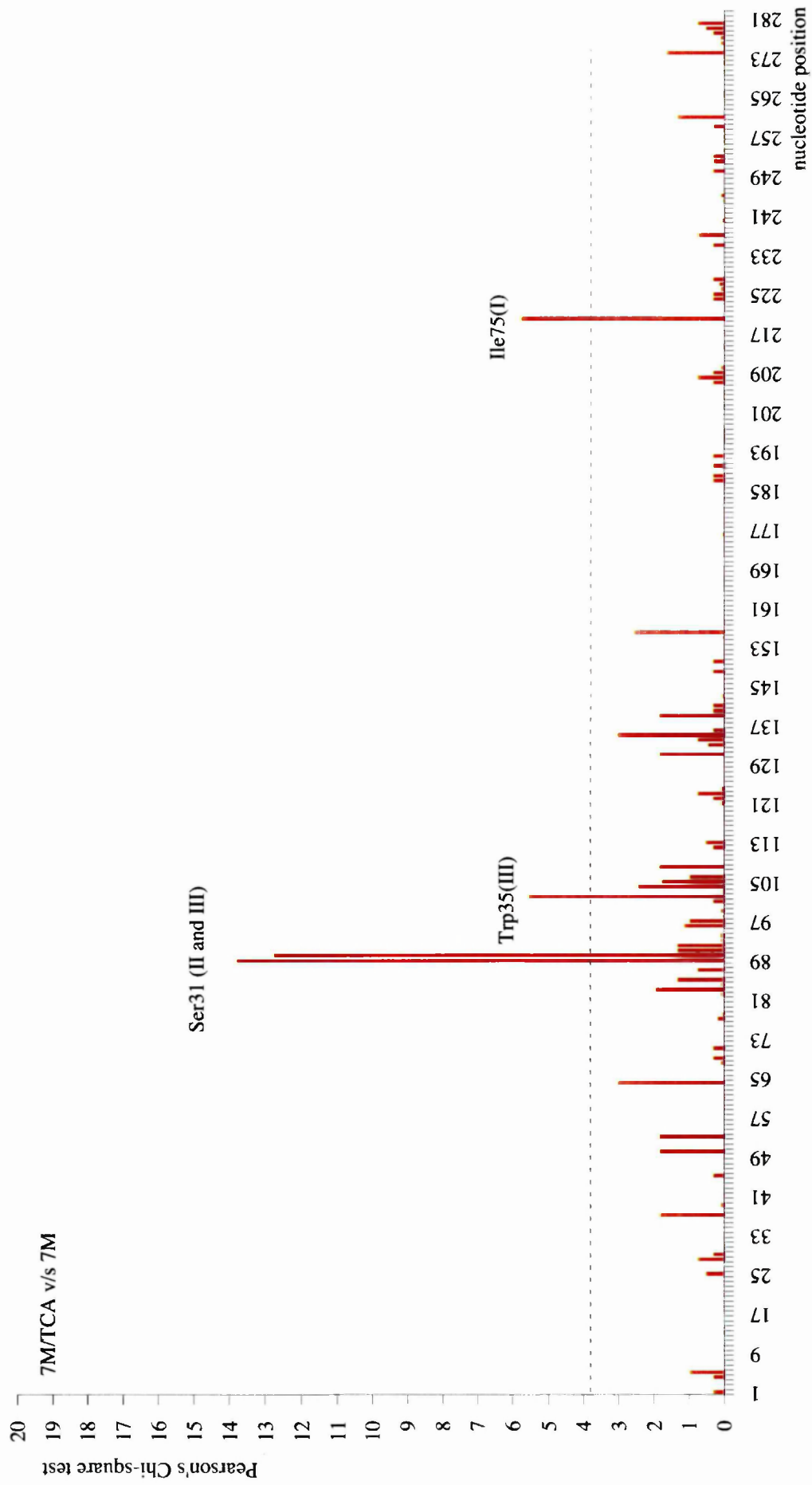
In addition, although with lesser significance, differences in the mutations in W35(III) and I75(I) were found. These position are more mutated in TCA than in 7M, but when TCA was compared with the WT they were not significant because these positions are hotspots in the large database. Therefore that reflect the necessity of large databases to avoid mistakes in the assessment.

4.4.3 'TA' motif as mutational target

The 'TA' motif is found hypermutated in Val30, Tyr36, Ser93 and Asn94. The change in the Tyr32 (TAC for TAT) introduced a new 'TA' core motif because the first base of Met33 (ATG) was mutated in 7% more 7M/TCA clones and in 4% more 7M clones compared to the WT. These differences were significant at the 5% level by the Chi-test for 7M/TCA v/s WT (6.5) and suggestive but not significant (2.7) for 7M v/s WT.

Figure 4.7. Pearson's Chi-square statistic test for the differences (increase/decrease) in the mutability resulting from local base substitutions in 7M/TCA v/s WT.

The test was performed for each position along the V κ gene in the wild type and the 7M/TCA mutant for the clones sequenced (239 and 61 respectively) with the number of mutations found. For significance at the 5% level, the Chi-square need to be greater than 3.84 (shown by a dashed line).



The difference between the 'TA' motif in both mutants could indicate that there are other elements like the palindromic structures, that may help in the targeting. Indeed, the DNA secondary structure around these residues is different between the two mutants (see Fig. 4.2). Tyr32 and Met33 codons in 7M/TCA can form a very small loop which may expose the 'TA' core for mutability. However, in the 'TA' core of the 7M mutant there is another loop more stable that competes.

4.5 The secondary structure has a lesser effect in the definition on hotspots

It is very difficult to assess the effect of hairpin loops independently of the effect of the DNA sequence motif on hotspots since in order to change the consensus sequences it is unavoidable to disturb the palindromic structures and vice versa. For example, the hairpin loop containing the Ser31 AGT is very stable in the CDR1 wild type. As G is at the tip of the loop, it is the most exposed base. This coincides with the fact that it is also the most mutated base of this triplet.

In the 7M mutant (Ser as AGC), the loop loses stability. However, both G and C are equally exposed at the end of the loop (see Fig. 4.2). The mutations in this case occur equally in both bases. On the other hand, the new TA core formed in the Tyr32-Met33 of the CDR1 mutants was more mutated when part of a putative loop. Therefore the palindromic structures may make the nucleotides accessible for hypermutation.

Figure 4.8. Pearson's Chi-square statistic test for the differences (increase/decrease) in the mutability resulting from local base substitutions in the 7M/TCA v/s 7M.

The test was performed for each position along the V κ gene in the 7M/TCA and the 7M mutants for the clones sequenced (61 and 75 respectively) with the number of mutations found. For significance at the 5% level, the Chi-square need to be greater than 3.84 (shown by a dashed line).

4.6 DNA sequence modifications and affinity maturation

Hybridomas were generated from the Lκ(WT&7M), LκNG and Lκ7M/TCA&Δ[Li]. The Vκ segment was amplified and cloned as described in Materials and Methods. The sequences are shown in the Appendix IV A, B and C respectively.

The analysis of these sequences revealed that the mutations introduced in the CDR1 did not disturb the introduction and selection of the characteristic aminoacid substitutions at His34 (by Gln or Asn) or Tyr36 (to Phe) observed in most of the antibodies of a secondary or tertiary immune response to phOx resulting in a ten-fold increase in affinity (Berek and Milstein, 1987). These mutations were often found in Lκ7M of Lκ(WT&7M) and Lκ7M/TCA hybridomas. They were found less frequently in the Lκ7M transgenes of LκNG hybridomas in spite of the change in His34 (CAC for CAT) (Appendix IV).

It was notable that, even within a single cell, there was a considerable difference in the number of accumulated mutations in individual transgene copies (Table 4.4). This phenomenon was well illustrated in the LκNG hybridomas, not only amongst Lκ and 7M but also in Lκ-Vgpt* and Lκ-V neo*Δ[XS]i transgenes. The latter exclude a skewing effect of antigen selection. Contrasting the mutations in e.g. C10.5.4 and C23.1.5 hybridomas, it is evident that this phenomenon is not inherent to the transgene itself as in one hybridoma one transgene accumulates more mutations and in the other hybridoma it is the other transgene.

Another intriguing observation was that the aminoacid substitutions characteristic for antibodies with increased affinity for phOx were not found randomly distributed in all the transgene copies. For example, both aminoacid substitution are present only in the 7M transgene and not in the WT transgene of the D3.3.3 and C12.1.3 hybridomas (which secrete only transgene light chain) derived from Lκ(WT&7M) line (Table 4.4). In the hybridomas form LκNG line the distinctive aminoacid substitutions of antibodies with increased affinity for phOx were found in 7 out of 9 hybridomas. Only one 7M transgene (in hybridoma B20.6) was found with one of the typical aminoacid substitution (Tyr36 for Phe). While in another (D1.5.5) with one of the substitution (Tyr36 for Phe) in a 7M copy also contained a WT copy with both, His34 and Tyr36 substitutions. The other 5 hybridomas had both aminoacid substitutions in the WT copies.

The finding of those mutations in specific copies may reflect differences in the transcription of those copies. In the Lκ(WT&7M) line the fact that one of the copies is truncated and does not contain 3' enhancer may be a reason for differential transcription. In the LκNG line there is evidence that some copies may not contain the 3' enhancer.

Table 4.4 Sequence analysis of the transgene V regions of the hybridomas from LκNG, Lκ(WT&7M) and Lκ7M/TCA&Δ[Li] transgenic mice.

LκNG	LκΔBWT		LκΔB1M		LκΔB7M		C	PCRe	LκVneo			LκVgpt	
C15.2.7	8(Q-F)	6	7	3	3	37	2	7	5	3	0	6	
C10.5.4	5(N-F)	8	6	1	1	23	9	9	7	3	0	1	
A6.8†	3	3	1	0		18	5	6	4	2	nf	3	
C23.1.5	7(N-F)	nf	nf	2	2	42	1	2	0			8	
D1.1.5	10(N-F)	0	nf	10(F)	0	25	3	3	1	0		0	
B20.6	4	0	2	4(F)	0	19	2	2	0			0	
C6.3.6	8(Q-F)	nf	4	5	2	14	1	0				0	
D12.3.8	0	0	2	0	0	29	2	0				1	
C4.6.8	0	0	2(F)	0	0	12	1	0				1	
D23.4	1	0	1	1	0	26	3	0				0	

Lκ(WT&7M)		LκWT	C	Lκ7M	C
D3.3.3§		3*	7	7(N-F)	5
C12.1.3§		5*	4	7(Q-F)	6
EH2.5		0	3	4	3
RB3.6		0	3	3	3
EA8.7		0	5	1	3
C323.2		0	2	0	5
C11.4.8.8		0	4	0	4

Lκ7M/TCA	Lκ7M/TCA	C	PCRe
B2.4.4	3	15	1
C14.1.9	6(N-F)	37	1
A20.5.2	4(N-F)	2	0

The table summarises the sequence results obtained from the LκNG, Lκ(WT&7M) and Lκ7M/TCA&Δ[Li] hybridomas. The result for LκVneo and Lκ-Vgpt (Yélamos et al., 1995) are included for comparison. The number of mutations is given for each transgenic copy. When no mutations are found in several M13 clones, they may derived from more than one unmutated copy.

C, number of clones found for each copy. Mutants are defined by a minimum of two identical sequences. For LκΔB (7M and WT) in LκNG a minimum of 15 clones were sequenced, but only the numbers of clones completely read is shown.

nf, the expected copy was not found among a minimum of 25 clones. Absent copies could result from instability of the transgenic DNA in B cells in the derived hybridomas.

PCRe, the number of mutations attributed to erroneous PCR amplifications, since each of them was observed in a single sequence only.

* These copies carry deletions which result in a coding sequence out of frame.

† The LκΔB transgene copies do not provide the light chain of the phOx antibody secreted by this hybridoma. The rest of the LκNG hybridomas secrete only transgenic rat κ.

§ In these Lκ(WT&7M) hybridomas the light chain of the phOx antibody secreted is only provided by the transgene. The rest of the hybridomas secrete both, transgene rat κ and endogenous mouse κ chains. The amino-acid substitutions characteristic of antibodies with increased affinity for phOx are indicated as H34 for N (N) or for Q (Q) and Y36 for F (F).

4.7 Discussion

All the mice used for the studies above contained multiple transgenes. This was particularly useful in the case of the LκNG line because the Lκ transgenes served as internal control (mainly for analysis of mutation rates) of the non immunoglobulin transgenes.

Vκ mutants with alterations of DNA sequence in the CDR1 were able to hypermutate at the same level and with the same intrinsic preferences as the wild type. This is in spite of the finding that one of the lines (Lκ(WT&7M) contains a truncated copy which lacks the 3' enhancer. Thus, it seems that the 3' enhancer of the intact neighbour copy is able to drive hypermutation independently of its downstream position with respect to the V-C elements. While these results indicate that the usual 3' enhancer position in the κ transgene is not required to achieve a high level of mutations, the hybridoma results suggest a slightly different picture (see below).

Some of the CDR1 modifications produced important changes in the well characterised hotspots of the VκOx1 CDR1. The second base of Ser31 AGT T is the major intrinsic hotspot within the mouse VκOx1 gene. It matches the Pu-G-Py-A/T consensus in which the most favoured base for mutations is the G. When Ser31 (AGT) was changed (for AGC) in the 7M, not only the second base but also the third is a hotspot. Ser26 is another intrinsic hotspots in the CDR1 of VκOx1 encoded by the AGCT consensus. Also in this case both bases, the G and the C, are mutated. Another example is the major hotspot found in the Vgpt gene, which match the GAG CTC consensus. The third base of Glu70 and the first base of Leu71 also mutate equally well (Yélamos et al., 1995).

The formation of the new hotspot in the third base of Ser31 indicates the importance of the primary DNA sequence for the targeting of specific nucleotides for mutation. This was further supported by the findings that Ser31 TCA was not a hotspot. A preference for the usage of AGY codons in the CDRs (particularly in the CDR1) and for TCNs in the frameworks have been observed in human V genes (Wagner et al., 1995) and also in *Xenopus* and sheep (Schwager et al., 1989; Reynaud et al., 1995). It suggests that the DNA sequence of germline V genes has evolved in response to selection for appropriately targeted mutability (Wagner et al., 1995). The results shown here indeed confirm the mutational machinery preferences for AGY encoded Ser and the bias against TCA encoded Ser as target for hypermutation. The 7M/TCA mutant demonstrates that the codon bias exists and that it correlates with the local nucleotide sequence

Not all the hotspots conform to the consensus above. For example, Val30(III), Tyr36(II), Ser93(I) and N94(I) are intrinsic hotspots that have a common 'TA' core with A being the mutated base. Indeed there are several examples amongst hotspots of λ chains (González-Fernández et al., 1994b). One of the modifications introduced in the CDR1 mutants was a change of Tyr32 (TAC for TAT). This generates a new 'TA' core motive as the first base of Met33 (ATG) was found mutated in 7M/TCA clones but not in the

7M ones. Further analysis is required to find a bigger consensus sequence, which includes the 'TA' core. No consensus has emerged from a preliminary analysis of the bases flanking the 'TA' core of the mutants.

On the other hand, these observations support the idea that hotspots are only partially created by local DNA sequence. Other structural features, such as palindromes, may well be important (González-Fernández et al., 1994b). The palindromic structures around Tyr32 and Met33 codons are different between the two mutants. Tyr32 and Met33 in the 7M/TCA mutant can form a very small loop which allows the 'TA' core to be more "exposed" while in the 7M the 'TA' core is also contained in another more stable loop. However Ser31 in the WT is located at the tip of a very stable hairpin loop that becomes unstable in the 7M mutant, yet it mutates even more in spite to the apparent loss of stability of the loop. I conclude that while the hairpin loops do not seem to be important to define which base is going to be targeted, they may have an enhancement role by making the 'favoured' nucleotides accessible for hypermutation.

In the hybridomas from the LκNG line a considerable difference in the number of accumulated mutations in individual transgene copies was observed in selectable and not selectable transgenes. This result confirmed the skewing previously noted using selectable transgenes (Rogerson et al., 1991) and demonstrated that it is not attributable to antigen selection.

On the other hand, the analysis of the aminoacid substitutions that increase antibody affinity in Lκ(WT&7M) and LκNG hybridomas suggested that those mutations occur more frequently in some transgenes than in others. e.g. in WT transgenes of LκNG line or in 7M transgenes of Lκ(WT&7M) line.

In Peyer's patch cells the differences in the frequency of mutations between the WT and 7M copies in both (Lκ(WT&7M) and LκNG) lines correspond to results for hybridomas. They are however not very significant. In addition, the hybridoma analysis is based on differences collected from a small database. The variation of the preferential targeting of the mutations associated with the affinity maturation in these copies suggests that it could arise from a difference in the efficiency of transcription of those copies. This could be confirmed with the analysis of the transcription levels of both (7M and WT) copies in several hybridomas from Lκ(WT&7M) line.

There are no extensive studies indicating that all the Lκ copies (without stop codons) are equally transcribed in the transgenic mice. Indeed some such differences have been noted in transgenes with premature stop codons in hybridomas (Lozano et al., 1993). It is also unknown whether tail-to-tail tandem integrations, which bring the two 3' enhancers relatively close, are better expressed than others.

5.1 Controlling elements in the J–C κ intron

Over the last decade our knowledge of somatic hypermutation has grown considerably. Little, however, is known about the molecular mechanism driving this process. The observation of hypermutation of rearranged κ transgenes in many different transgenic lines suggests that hypermutation depends on *cis*-acting elements in the κ locus and is largely independent of chromosomal position.

It has been shown that multiple *cis*-acting DNA elements are involved in the regulation of somatic hypermutation (Betz et al., 1994). Both, the J–C κ intron enhancer/MAR region and the 3' enhancer downstream of the C κ region are required for full hypermutation. The requirement for these elements suggests that mutation correlates with an "open" DNA structure associated with the potential for transcription. Hypermutation seems to occur only in germinal centre centroblasts, which express no surface IgD. There is, however, no evidence which directly connects hypermutation to the immunoglobulin transcription rate in these cells. The role of the enhancers and transcription in hypermutation remains unclear.

The J–C κ intron contains a matrix attachment region (MAR) (Cockerill and Gallard, 1986) and the κ intron enhancer (Ei) (Queen and Baltimore, 1983; Picard and Schaffner, 1984a). To delineate *cis*-acting DNA elements within the J–C intron required to target hypermutation two constructs were designed to generate transgenic mice harbouring independent deletions of intron enhancer (L κ Δ B Δ [Ei]) and matrix attachment region (L κ Δ B Δ [MAR]) (Fig. 5.1). Detailed description of these constructs can be found in Materials and Methods.

L κ Δ B Δ [Ei] carries the deletion of the region which contains the κ A, κ B, E1, E2, E3 sites. These sites were characterized both functionally and in binding studies and were shown to be crucial for the activity of this enhancer (Lenardo et al., 1987; Nelms et al., 1990). Other less well characterized sites, including a silencer region (Pierce et al., 1991), a non consensus octamer binding site (Currie and Roeder, 1989) and the κ NE and κ BS sites (Sak sela and Baltimore, 1993) are also included in the deleted area of L κ Δ B Δ [Ei].

The other mayor DNA element in the J–C κ intron, the matrix attachment region was deleted in L κ Δ B Δ [MAR]. MARs are 70% A–T rich sequences that mediate the association of the nuclear DNA with the nuclear matrix (scaffold) and they are

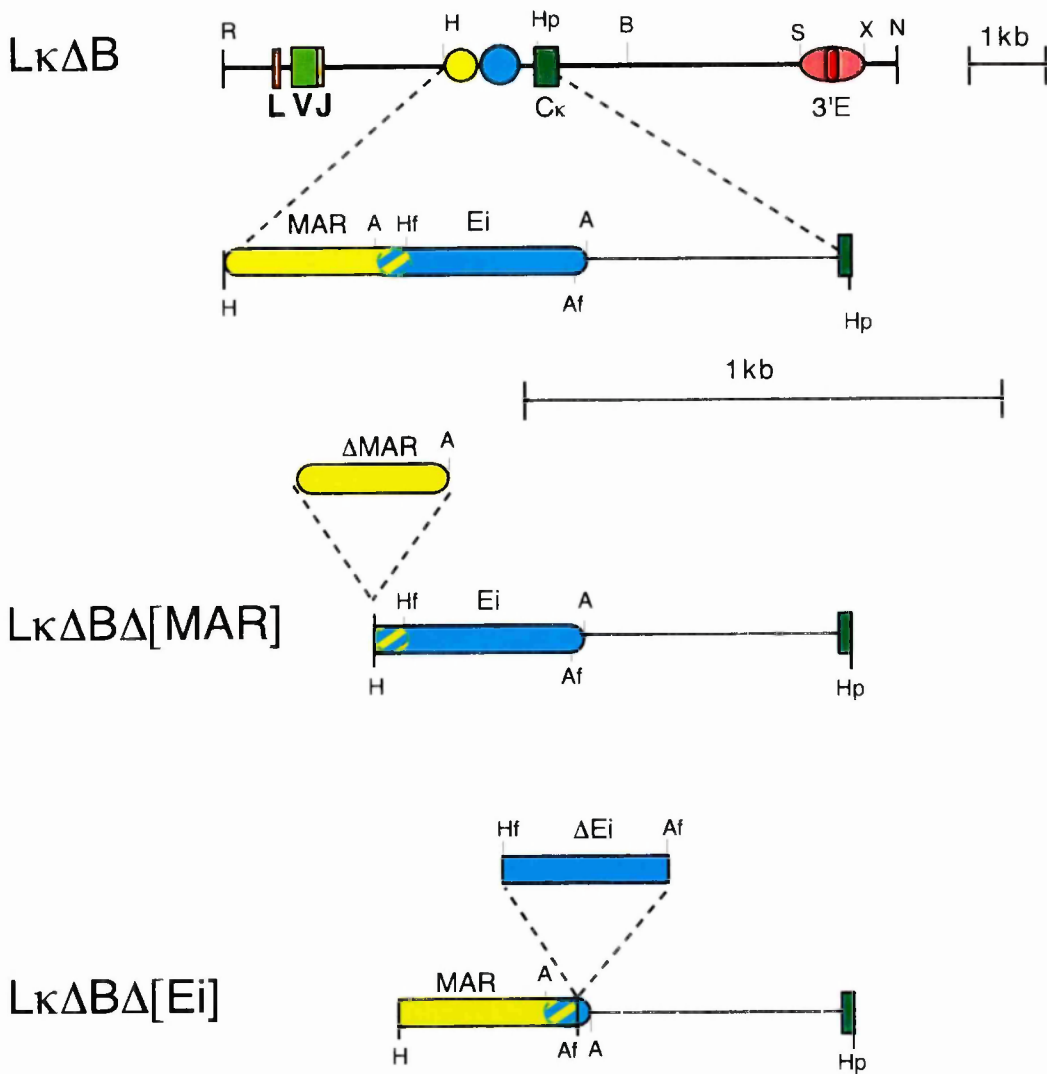


Figure 5.1 a Structure of transgenes. The constructs are modified versions of L_{κ} (Sharpe et al., 1991). The MAR (HindIII-HinfII) and the κ intron enhancer (AluI-AluI) regions in the $L_{\kappa\Delta B}$ are shown enlarged for reference. The constructs containing the deletions are all analogous to $L_{\kappa\Delta B}$, differing only in the region between the HindIII and HpaI sites, as shown in the enlargements. R, EcoRI; H, HindIII; Hp, HpaI; B, BamHI; N, NotI; A, AluI; Hf, HinfI; Af, AflIII.

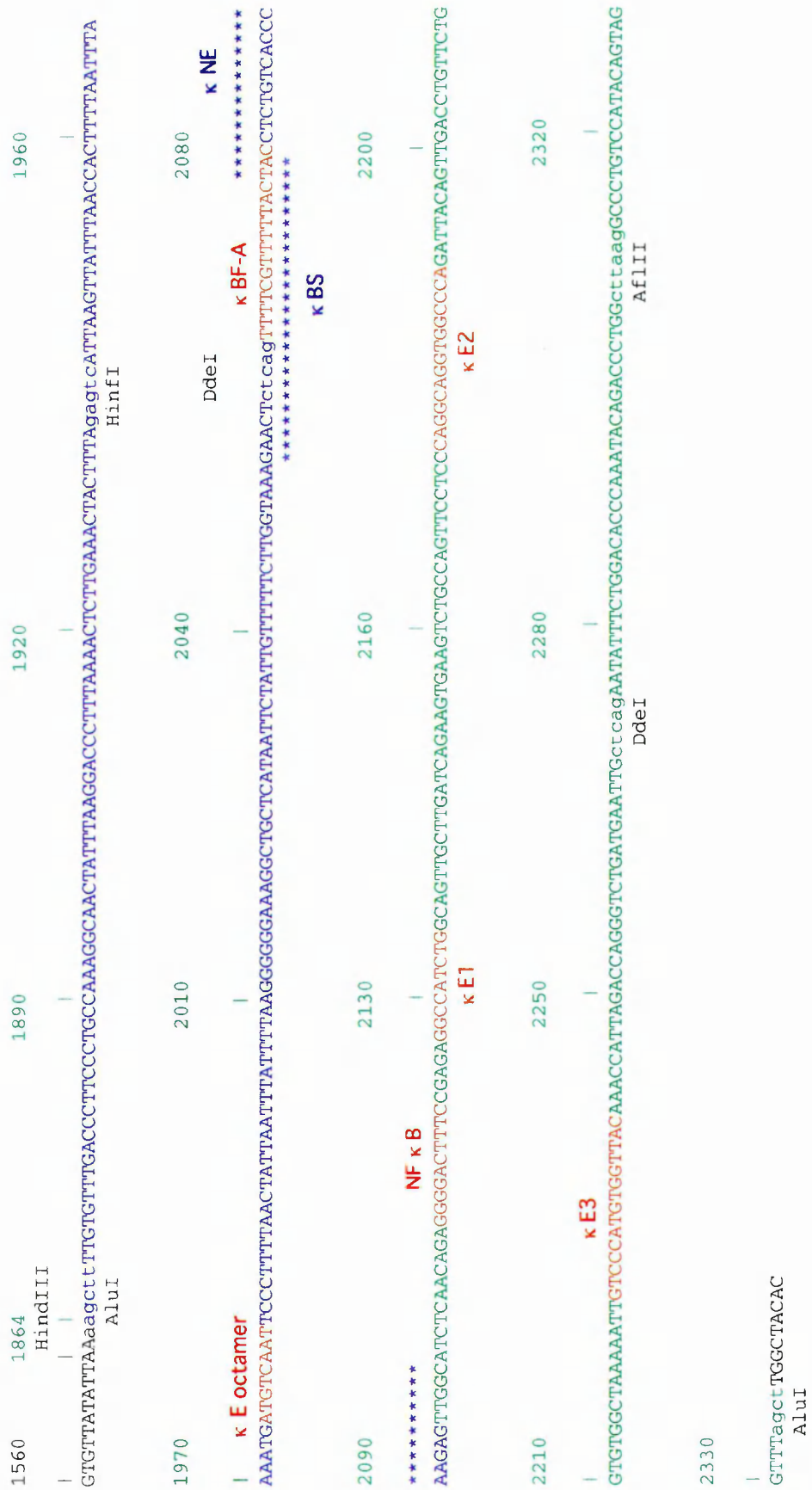


Figure 5.1 c J-C κ intron relevant DNA sequence in the L κ Δ B[Δ MAR] construct which carries the deletion of the MAR. The deleted fragment (1573-1863) includes most of the DNA sequence of the MAR (see Fig. 5.1.b). The Ei DNA fragment (AluI[1864]-AluI[2334]) is highlighted in green. The core enhancer is located between the two DdeI sites (Fulton and Van Ness, 1994). The negative regulatory elements are indicated by asterisks. The 232bp silencer is highlighted in blue (Pierce et al., 1989). Other sequence motives are indicated in red. Some restriction sites are shown in lowercase for reference.

preferentially bound and cleaved by Topoisomerase II. Several DNA motifs have been found in MARs of different species. These motives are repeated several times along the MAR and are also found in origins of replication, and in the recognition site of several homeodomain proteins. The sequences are: AATATTTT or subsets thereof (Cockerill and Garrard, 1986); TT/TTT/ATTT/ATT, and stretches of ATATTT -together with its variants, ATATTTT and ATATT-, ATTA, ATTTA (Boulikas, 1993).

As MARs are defined by physical rather than functional characteristics, their functions (other than to anchor chromatin to the nuclear matrix) may be heterogeneous. MARs have been associated with transcription (Blazquez et al., 1989; Stief et al., 1989; Forrester et al., 1994), replication (Cockerill and Garrard, 1986; Dijkwel and Hamlin, 1988), recombination (Sperry et al., 1989), demethylation of the κ gene (Lichtenstein et al., 1994), position-independent expression to heterologous genes (MacKnight et al., 1992) and regulation of the cell type specificity of the Ig H intron enhancer (Scheuermann and Chen, 1989; Zong and Scheuermann, 1995). Its potential role in somatic hypermutation is analysed in this Chapter.

5.1.1 Hypermutation of Δ Ei κ transgenes

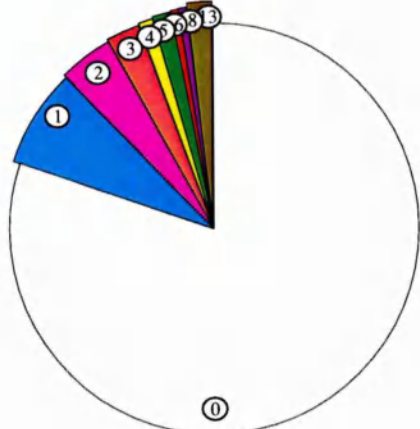
Hypermutation of Δ Ei transgenes was analysed by polymerase chain reaction (PCR) amplification from a population of germinal centre cells from Peyer's patches with the oligonucleotides VKOXFOR and VKOXBACK (Appendix II) using Pfu polymerase. PCR products were cloned into M13mp18 as described in Materials and Methods.

Sequences from two mice were analysed: the founder Δ Ei 9628 (5 months and 26 days old) and one of its offspring Δ Ei 9701 (4 months and 14 days old). The sequencing primer was the oligonucleotide JK5. The PCR error frequency (0.13×10^{-3} per base) was determined by sequencing transgenes from hybridomas obtained from the L κ 7M/TCA Δ [Li] transgenic line (see Chapter 4).

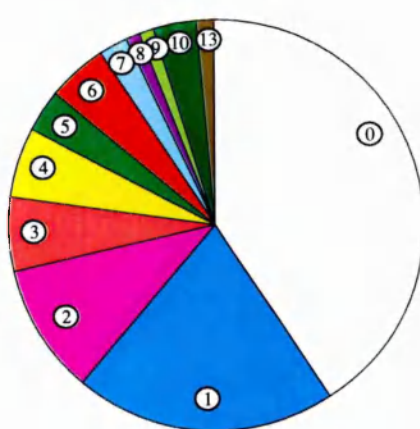
The sequence analysis of Δ Ei V κ regions revealed that the transgene had been a substrate for hypermutation. The sequences with one or more mutations are shown in the Appendix III A. Almost 20% of the 164 sequences were mutated with 12.2% carrying more than one mutation (Table 5.1 and Fig. 5.2).

The mutation frequency of all mutated clones and amongst those clones carrying more than one mutation are the usual quantitative values employed to compare the levels of hypermutation achieved by the transgenes. In order to have an additional indicator the number of mutations per clone (of all mutated clones and above and those carrying 2 mutation) was also included.

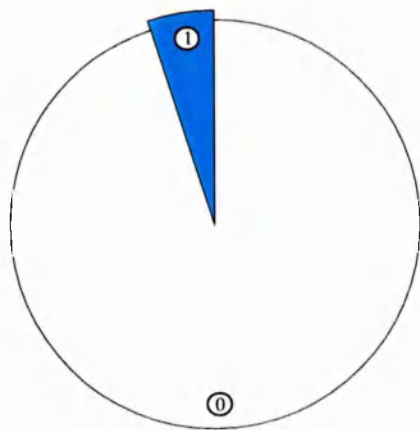
The mutation frequency based on the total number of clones (M_T) was corrected for errors introduced during the PCR amplification by subtracting directly the PCR error rate from the mutation frequency. This allows the comparison with mutation frequencies



$L\kappa\Delta[Ei]$



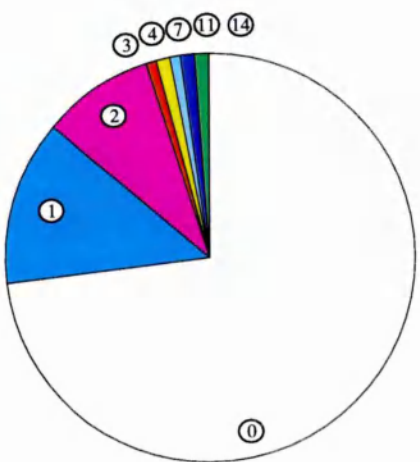
$L\kappa(WT)in L\kappa(WT\&7M)$



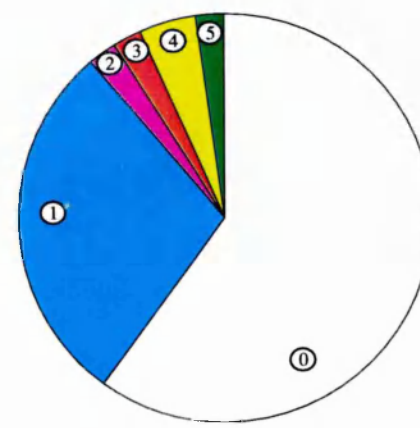
$L\kappa\Delta[MAR]$



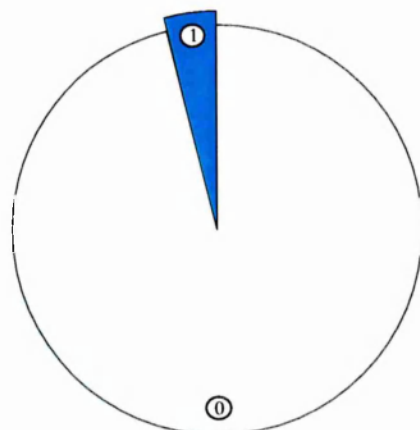
$L\kappa\Delta B(WT) in L\kappa NG$



$L\kappa\Delta[F13'EC]$

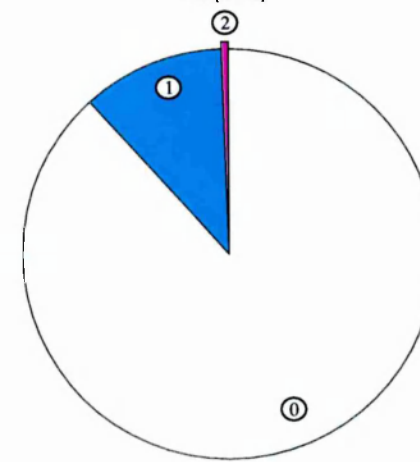


$L\kappa\Delta[3'EC]$



PCR error

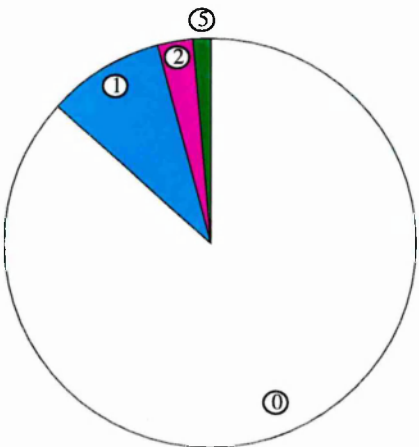
A



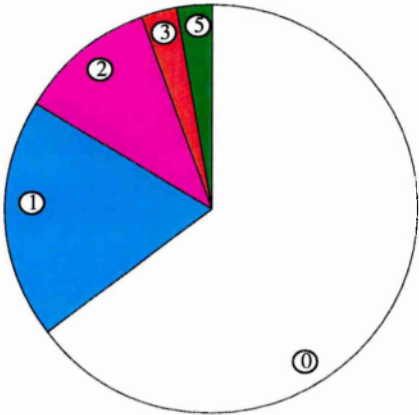
PCR error

B

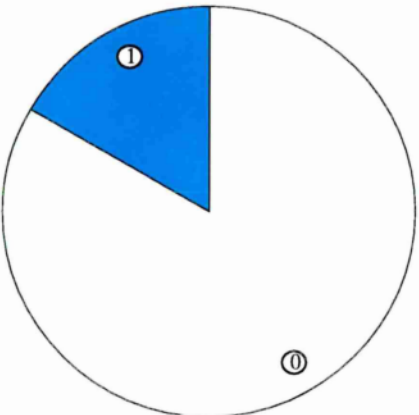
Figure 5.2 Mutations in the transgenic V region. Each pie chart depicts the proportion of sequences with the specified number of mutations. All the experiments performed with the same polymerase are grouped in columns. A pie chart showing the control for PCR errors of each set of experiment is included for comparison. A) PCR with Pfu polymerase. PCR error measured by sequencing Lκ7M/TCA&[Li] transgenes from characterized hybridomas. B) PCR with Taq polymerase. PCR error measured by sequencing LκNG transgenes from characterized hybridomas. C) The results for LκΔ[Ei/MAR] and LκΔ[E3'] (Betz et al., 1994) are included for comparison. The result for Lκ[E3'] (PNA low B cells) has the same value as the estimated PCR error (0.7×10^{-3}).



LκΔ[Ei/MAR]



LκΔ[E3']



LκΔ[E3'] (PNA low B cells)

obtained for other transgenes independent of the PCR amplification conditions used. The data from the LκΔ[Ei/MAR] transgene (Betz et al., 1994), which carries the deletion of both the intron enhancer and the matrix attachment region, are also included for comparison in Table 5.1.

Table 5.1 ΔEi and ΔMAR transgene mutations in PCR clones derived from Peyer's patch B cells (PNA^{hi}).

Transgene	Clones [#]			Mutations [*]		Mut./clone [□]		Mutations/10 ³ base pair [¶]	
	Total	=1	>1	Total	≥1	>1	MF [†]	MF(>1)	
ΔEi	164	12	20	109	3.4	4.8	2.3 (2.2)	17.2	
ΔMAR	42	2	0	2	1	0	0.16 (0.03)	-	
PCR error [‡]	52	2	0	2	1	0	0.1	-	
Lκ(WT) [§]	147	28	67	381	4	5.2	9.2 (8.8)	18.7	
PCR error [∅]	245	27	1	29	0.9	2	0.4	7	
LκΔ[Ei/MAR] [◇] 75	7	3	16	1.6	3	0.76 (0.06)	10.7		
PCR error [‡]							0.7	7	

[#] Total number of PCR-generated clones and the number of those carrying one (=1) or more than one (>1) mutations.

^{*} Sum of mutations in all clones.

[□] The number of mutations per clone was calculated using either the number of clones carrying mutations (≥1) or only those carrying more than one mutation (>1).

[¶] The frequency of mutations per 10³ base pairs was calculated using either the total number of clones (MF[†]) or only those clones carrying more than one mutation (MF(>1)). The values after correction of the PCR error are given in parentheses.

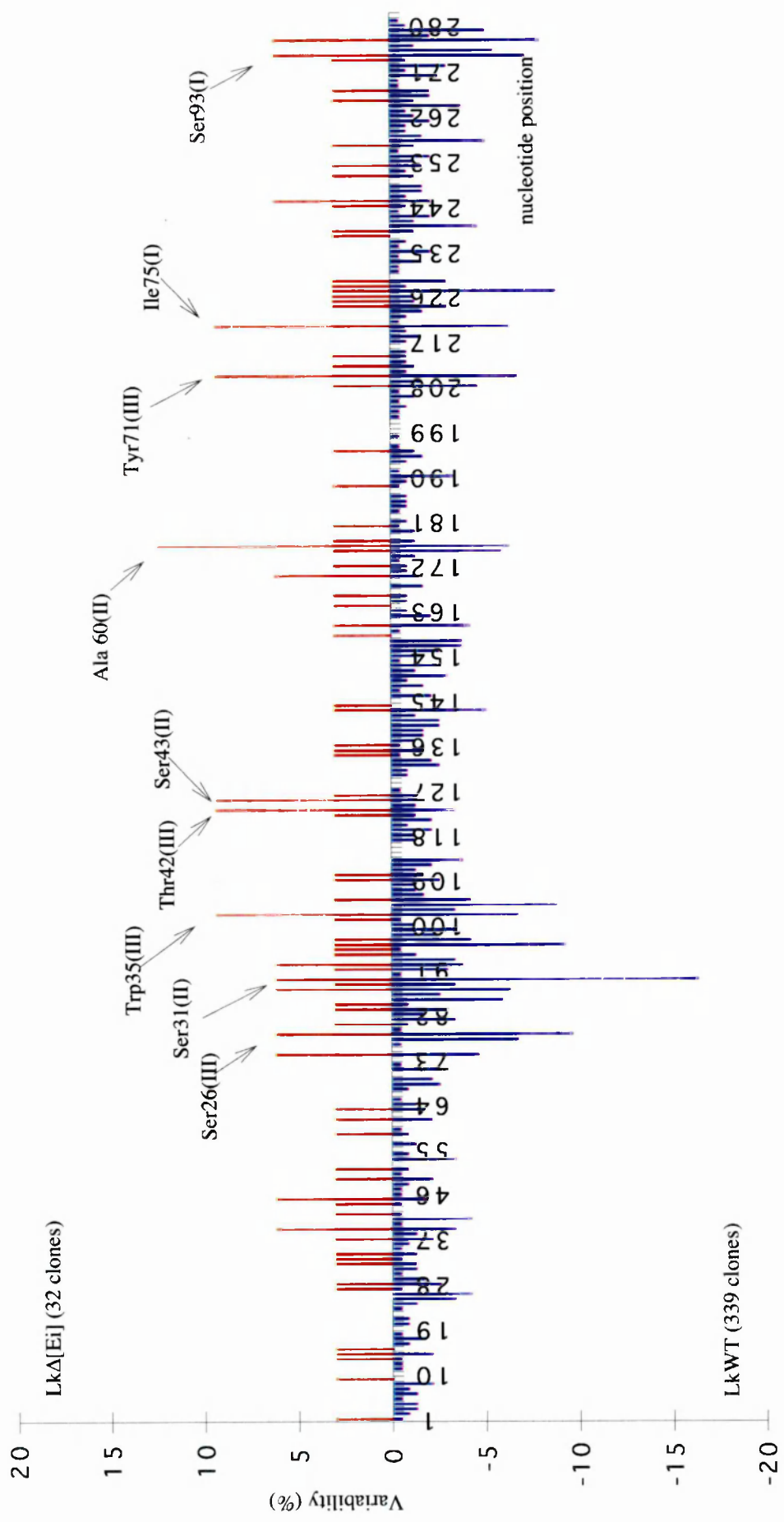
[◇] The LκΔ[Ei/MAR] data originate from Betz et al., 1994. The number of nucleotides sequenced was 279/clone.

[‡] The PCR error in this experiment was theoretically estimated. The MF(>1) for the PCR errors (7) was calculated for a minimum of one clone with 2 mutations.

[†] Clones obtained from hybridomas of Lκ7M/TCA&Δ[Li] transgenic line.

[§] Clones obtained from LκNG and Lκ(WT&7M) transgenic lines.

[∅] Clones obtained from hybridomas of LκNG transgenic line.



The mutation frequency amongst ΔEi clones carrying more than one mutation was 17.2×10^{-3} mutations per base pair. This value is comparable to the frequency of mutations observed in L κ transgenes (18.7×10^{-3} mutations per base).

The overall mutation frequency was 2.35×10^{-3} mutations per base pair, a value 4 times lower than the value obtained for L κ (WT) and 36 times higher than the value observed for L $\kappa\Delta$ [Ei/MAR]. This suggests that the removal of the κ intron enhancer does not result in the dramatic fall in hypermutation of the L $\kappa\Delta$ [Ei/MAR].

This low overall mutation frequency is due to an increased proportion of unmutated clones rather than to a decreased number of highly mutated clones.

This line has more than 6 copies of the transgene inserted in different places of the mouse genome, and some copies may not be able to mutate. This has been observed previously in the L κ NG line, where one of the four copies of the Vneo transgene did not mutate (Yélamos et al., 1995). There is another example where the number of non mutated V κ sequences found is large for unknown reasons. In the line L κ 7M/TCA& Δ [Li], which carries at least 1 copy of L κ 7M/TCA, 65% of the sequences analysed were found unmutated (see Chapter 4). This value is however lower than that obtained with L $\kappa\Delta$ Ei (80%).

In spite of the low frequency of L $\kappa\Delta$ Ei V κ mutated clones, the majority of these carry more than one mutation (see Fig. 5.2). Indeed the accumulated mutations amongst multiple mutated clones in L $\kappa\Delta$ Ei (4.8) and L κ (WT) (5.2) is similar, although there are fewer highly mutated clones in L $\kappa\Delta$ Ei. These results show that hypermutation is not impeded when the Ei is absent as it is in L $\kappa\Delta$ [Ei/MAR] where the number of mutations per clone (≥ 1) is only 1.6.

The distribution of somatic mutations along the L $\kappa\Delta$ [Ei] transgene is shown in Fig. 5.3, where the variability is given as the percentage of sequences that carry a mutation at each particular nucleotide of the V κ region. The segment analysed includes 282 bp using as the starting point the first coding nucleotide for the mature variable region.

Figure 5.3 Distribution of mutations in the L $\kappa\Delta$ [Ei] transgene (opposite page). The distribution of mutations amongst the transgenic V regions clones isolated from Peyer's patch PNA^{hi} B cells from the L $\kappa\Delta$ [Ei] transgenic line is shown. Variability at each nucleotide position is computed as the percentage of sequences carrying one or more mutations which are mutated at that position. Hotspots are indicated. The distribution of mutations on L κ (WT) transgenes is provided for comparison.

A large compilation of transgenic L κ (WT) sequences from Peyer's patch B cells (see Chapter 4) is used for comparison to assess whether the distribution of mutation is affected when the Ei was deleted.

The compilation of mutations in the L κ Δ Ei transgene shows the expected non random distribution of mutations and the presence of hotspots preferentially around CDR1. Although the number of mutations is relatively small, it is possible to find some of the most dominant intrinsic hotspots previously identified in the large data base for the same gene: the third base of Ser-26, the second base of Ser-31 and the second base of Ala-60 (Betz et al., 1993b)). There are other apparent hotspots (the third base of Trp-35, the third base of Thr-42, the second base of Ser-43, the third base of Tyr-71 and the first base of Ile-75) that are less prominent in the large L κ database (Fig. 5.3).

The nucleotide substitution preferences of mutations in L κ Δ Ei Peyer's patches (Table 5.2) is comparable to the previously defined substitution data from L κ 6 Peyer's patches (González-Fernández and Milstein, 1993).

Table 5.2 Nucleotide substitution in mutated L κ Δ Ei transgenes from Peyer's patch PNA^{hi} B cells.

From	To				Total mutations*
	T	C	A	G	
T	-	10 (0.59)	3 (0.18)	4 (0.23)	17 (0.17)
C	27 (0.73)	-	4 (0.11)	6 (0.16)	37 (0.30)
A	10 (0.34)	3 (0.10)	-	16 (0.55)	29 (0.27)
G	4 (0.15)	5 (0.19)	17 (0.65)	-	26 (0.26)

The proportion of mutations in each base corrected for the base composition of the region analysed is shown in parentheses.

* Number of mutations obtained for each of the four bases.

The predominant substitutions show the bias of transitions (70 changes out of 109) over transversions (39 changes out of 109). The bias against mutations in pyrimidines, which indicates the polarity of the hypermutation machinery, is unambiguous only for T, since C mutated more than G and A. This discrepancy with previously reported data is likely to be due to the small number of mutated clones in the L κ Δ Ei database.

5.1.2 Hypermutation of Δ MAR κ transgenes

Germinal centre PNA^{hi} B cells from Peyer's patches were collected from one of the descendants (9583) (4 months and 25 days old) to analyse hypermutation as described in Materials and Methods. The V κ O \times 1 gene segment was amplified with Pfu polymerase using the oligonucleotides BEALEADER and VKOXFOR (Appendix II). For sequencing the primer JK5 was used. The PCR error frequency was the same as described in section 5.1.1. The sequences with one or more point mutation are shown in the Appendix III B.

The sequence analysis of the Δ MAR V κ region revealed that 95.4% of the sequences are unmutated (see Table 5.1 and Fig. 5.2). There were no clones carrying more than one mutation among 42 clones sequenced. The Δ MAR transgene was therefore not a substrate for hypermutation. The low overall mutation frequency (MFt) (0.16×10^{-3} per base pair) is the same as the error frequency of the PCR procedure itself and it is more than 100 times lower than the mutation frequency for the L κ (WT).

These results are even more extreme than those previously obtained for the L $\kappa\Delta$ [Ei/MAR] transgenic line carrying the deletion of both, MAR and Ei (Betz et al., 1994). In both cases the MFt is very low, in the range of the value obtained for the PCR error. The MF amongst the Δ [Ei/MAR] mutated clones was 10.7×10^{-3} per base pair, but it should be noticed that it is less than two times higher than the minimum theoretical value (7×10^{-3} per base pair). However, there were three clones carrying more than one mutation in the L $\kappa\Delta$ [Ei/MAR] but none in L $\kappa\Delta$ [MAR]. This indicates that even when hypermutation is severely affected it is still possible to find some mutated clones.

The hypermutation analysis of Δ MAR transgenes indicates a crucial role of MAR on this process.

In both (L $\kappa\Delta$ [Ei] and L $\kappa\Delta$ [MAR]) transgenic lines a number of clones with deletions were found. The frequency of deletion in L $\kappa\Delta$ [Ei] and L $\kappa\Delta$ [MAR] was compared with that in other transgenic lines (Table 5.3). The frequency of deletions observed in L $\kappa\Delta$ [Ei] and L $\kappa\Delta$ [MAR] clones are not significantly higher than in other transgenic lines. None of the clones sequenced carried insertions.

Table 5.3 Transgene deletions in PCR clones derived from Peyer's patch B cell (PNA^{hi}) in different transgenic lines.

Transgenic line	Clones with deletion [†]		Clones analysed*	Frequency of deletions (per 10 ⁴ bases) [‡]
	1bp	>1bp		
LκΔ[Ei]	1	4	164	1.1
LκΔ[MAR]	2	0	42	1.7
Lκ7M/TCA	2	1	172	0.6
Lκ in Lκ(WT&7M)	0	1	88	0.4
Lκ7M in Lκ(WT&7M)	1	0	81	0.4
LκΔB7M in LκNG	3	0	26	4.1
LκΔB in LκNG	1	1	59	1.2
LκΔ[3'EC]	0	0	45	0
LκΔ[F13'EC]	0	0	101	0
Lκ-Vgpt [§]	1	0	67	0.5
Lκ-VneoΔ[XS]i [§]	0	0	57	0
Lκ6 [¶]	1	0	222	0.1
Lκ3 [¶]	0	1	76	0.5
Lκ[3'F] [¶]	2	0	65	1.1
LκΔ[E3'] [¶]	0	0	37	0
LκΔ[Ei/MAR] [¶]	0	0	75	0
LκΔB in ENG [¶]	0	1	20	1.7
Lκ[p5] ^{βG#}	0	0	96	0
Lis 1718 [#]	0	0	94	0
Lis 1922	1	0	141	0.2

[†] The number of clones carrying deletions (either 1bp or more than 1bp) is indicated. The number of clones carrying deletions found in transgenic lines that I have not analysed is included for comparison.

* Total number of clones sequenced.

[‡] Frequency of deletions is given as the number of deletions found per 10⁴ bp sequenced.

[§] J. Yélamos, personal communication.

[¶] A. González-Fernández, personal communication.

[#] C. Rada, personal communication.

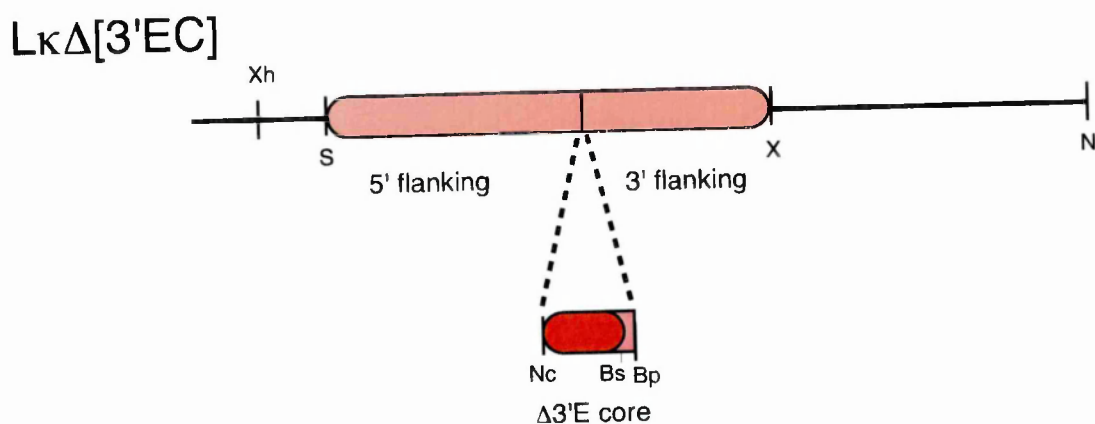
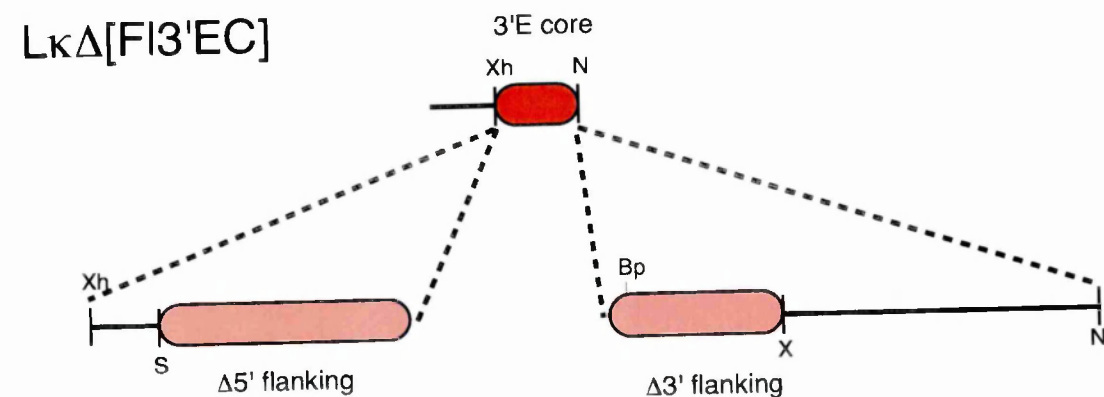
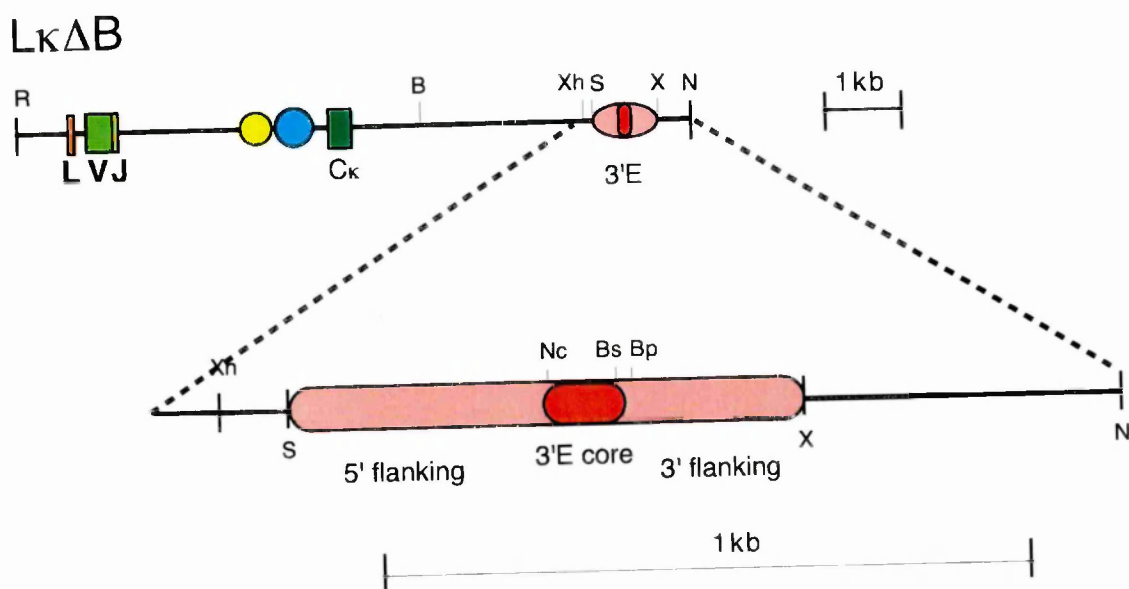


Figure 5.4 a Structure of transgenes. The constructs are modified versions of L κ (Sharpe et al., 1991). The XhoI-NotI fragment [containing the 3' enhancer (SacI-XbaI) region and its core (NcoI-BstXI)] in the L $\kappa\Delta$ Bam are shown enlarged for reference. The constructs, containing the deletions of the core enhancer or its flanking regions, are analogous to L $\kappa\Delta$ Bam. They differ only in the region between the XhoI and NotI sites as shown in the enlargements. R, EcoRI; H, HindIII; Hp, HpaI; B, BamHI; N, NotI; X, XbaI; Xh, XhoI; S, SacI; Nc, NcoI; Bs, BstXI; Bp, BspMI.

5.2 Controlling elements in the κ 3' enhancer

The 3' enhancer is one of the elements that regulate the cell specific and developmental expression of the κ light chain. This enhancer was found 8.5kb downstream of the constant region exon and is stronger than the κ intron enhancer (Meyer and Neuberger, 1989).

Transient transfection assays using reporter genes revealed that most (75%) of the enhancer activity is localised in a 132 bp core (Meyer et al., 1990; Pongubala and Atchison, 1991) (Fig. 5.4). However the core enhancer only accounted for 30-50% of the activity when assayed in stable transfected S107 cells with the L κ and L $\kappa\Delta$ B constructs carrying the core instead of the whole 3' enhancer (see Chapter 3).

The involvement of HLH proteins in the regulation of this enhancer was substantiated by repression of the enhancer through the HLH-inhibitory protein Id (Benezra et al., 1990; Pongubala and Atchison, 1991). Transient transfection assays with the core enhancer revealed that it contains at least three functional DNA sequences (PU.1/NF-EM5, E2A and κ E3'-CRE) (Pongubala and Atchison, 1991; Pongubala et al., 1992; Pongubala et al., 1993; Pongubala and Atchison, 1995). However, the flanking regions of the core enhancer seem to be involved in the developmental control of the 3' enhancer activity (Pongubala and Atchison, 1991; Meyer and Ireland, 1994).

The same DNA region that contains the 3' enhancer was shown to be essential for hypermutation and a link between transcription of the immunoglobulin gene and the hypermutation machinery was suggested (Betz et al., 1994; Milstein and Rada, 1995; Neuberger M. S. and Milstein, C., 1995; Peters and Storb, 1996). The L κ construct, which encodes an immunoglobulin κ light chain and includes 9kb of DNA downstream of the C κ , was the first κ transgene shown to undergo hypermutation at the same level as in the endogenous genes (Sharpe et al., 1991). In contrast the transgenes used in previous studies (O'Brien et al., 1987, Sharpe et al., 1990), mutated only at a very low level. These results suggested that the 3' flanking region of the κ locus contained control elements for both hypermutation and transcription. This was confirmed when the κ 3' enhancer was deleted resulting in a large reduction of somatic mutation (Betz et al., 1994).

None of these studies showed whether the 3' enhancer core includes the necessary elements to target somatic hypermutation. To extend the systematic delineation of the *cis*- acting DNA elements within the 3' enhancer, which are involved in control of hypermutation, two transgenic lines harbouring independent deletions of the enhancer core (L $\kappa\Delta$ [3'EC]) or the flanking regions of the 3' enhancer core (L κ [F13'EC]) (Fig. 5.4) were made. Detailed description of these constructs can be found in Materials and Methods.

5.2 1 Hypermutation of LκΔ[F13'EC] transgenes

Hypermutation in the Vκ region of the LκΔ[F13'EC] transgenes, which lack the flanking regions around the 3' enhancer core, was analysed in germinal centre cells of Peyer's patches by amplification with the oligonucleotides VKOXFOR and VKOXBACK using Pfu polymerase. PCR products were cloned in M13mp18. The oligonucleotide JK5 was used as the sequencing primer. Sequences from two transgenic mice were analysed: the Δ[F13'EC] 4256 (5 months and 13 days old) and Δ[F13'EC] 4272 (4 months and 15 days old). The PCR error frequency (0.13×10^{-3} per base) was determined by sequencing transgenes from hybridomas obtained from the Lκ7M/TCA&Δ[Li] transgenic line (Chapter 4).

The sequence analysis of the Δ[F13'EC] Vκ regions revealed that the transgene had been a substrate for hypermutation. The sequences with one or more point mutations are shown in the Appendix III. C. Almost 27% of the 101 sequences were mutated and 14% carried more than one mutation (Table 5.3 and Fig. 5.2).

The mutation frequency based on the total number of clones (Mft) was corrected for errors introduced during the PCR amplification by subtracting directly the value obtained for the PCR errors (Table 5.4) to allow the comparison with mutation frequencies obtained for other transgenic lines independent of the PCR conditions used in their amplification. The data are compared with those of the LκΔ[E3'] transgene, where the whole 3' enhancer was deleted (Betz et al., 1994).

The mutation frequency amongst clones carrying two or more mutations was 14.4×10^{-3} per base pair. This value is lower but comparable to that observed in Lκ (WT) transgenes (18.7×10^{-3} mutations per base) but higher than that obtained in LκΔ[E3'] (9.5×10^{-3} mutations per base).

The corrected overall mutation frequency (2.3) is more than ten times higher than the PCR error, indicating that the removal of the flanking regions around the 3' enhancer core does not result in the drastic decrease of hypermutation rate found with LκΔ[Ei/MAR] and LκΔMAR (see previous section). It is however 3.8 times lower than the value obtained for Lκ (WT) transgenes.

This low overall mutation frequency could be due to an increased proportion of unmutated clones that originate from truncated transgene copies not able to mutate. In this line only 2 copies contain the 3'E fragment, which leaves about 4 Vκ fragments which are possibly unable to mutate (for details see Chapter 3). Indeed only 27% of all LκΔ[F13'EC] transgene clones sequenced (Fig. 5.2) were mutated while 52% of the mutated clones carried more than one mutation. Several clones were highly mutated (up to 14 mutations) and the population of clones carrying more than one mutation contained an average of 4.1 mutations per clone. This value is comparable to that obtained in the Lκ(WT) (5.2). In LκΔ[E3'] transgene, where the full 3' enhancer was deleted, only one

clone with 5 mutations (also the highest mutated) was found. The number of mutations per clone (in clones carrying more than one mutation) was 2.7. These results suggest that the 3'E flanking regions are much less important for somatic mutation than the whole of the 3' enhancer.

Table 5.4 LκΔ[F13'EC] and LκΔ[3'EC] transgene mutations in PCR clones derived from Peyer's patch B cells (PNA^{hi}).

Transgene	Clones [#]			Mutations [*]		Mut./clone [□]		Mutations/10 ³ base pair [¶]	
	Total	=1	>1	Total		≥1	>1	MFt	MF(>1)
LκΔ[F13'EC]	101	13	14	70		2.6	4.1	2.4 (2.3)	14.4
PCR error [†]	52	2	0	2		1	0	0.1	0
LκΔ[3'EC]	45	13	5	31		1.7	3.6	2.4 (2)	12.8
WT [‡]	147	28	67	381		4	5.2	9.2 (8.8)	18.7
PCR error [∅]	245	27	1	29		0.9	2	0.4	7
LκΔ[3'E] [◇]	37	7	6	23		1.7	2.7	2.2 (1.5)	9.5
PCR error [‡]								0.7	7

[#] Total number of PCR-generated clones and the number of those carrying one (=1) or more than one (>1) mutations.

^{*} Sum of mutations in all clones.

[□] The number of mutation per clone was calculated using either the number of clones carrying mutations (≥1) or only those carrying more than one mutation (>1).

[¶] The frequency of mutations per 10³ base pairs was calculated using either the total number of clones (MFt) or only those clones carrying two or more mutations (MF(>1)). The values after correction of the PCR error are given in parenthesis.

[◇] The LκΔ[3'E] data was taken from Betz et al., 1994 The number of nucleotides sequenced was 279/clone.

[‡]The PCR error in this experiment was theoretically estimated. The MF(>1) for the PCR errors was calculated for a minimum of one clone with 2 mutations.

[†] Clones obtained from hybridomas of Lκ7M/TCA&Δ[Li] transgenic line.

[‡] Average obtained from LκΔB and Lκ clones from LκNG and Lκ(WT&7M) transgenic lines (Chapter 4).

[∅] Clones obtained from hybridomas of LκNG transgenic line.

5.2.2 Hypermutation of LκΔ[3'EC] transgenes

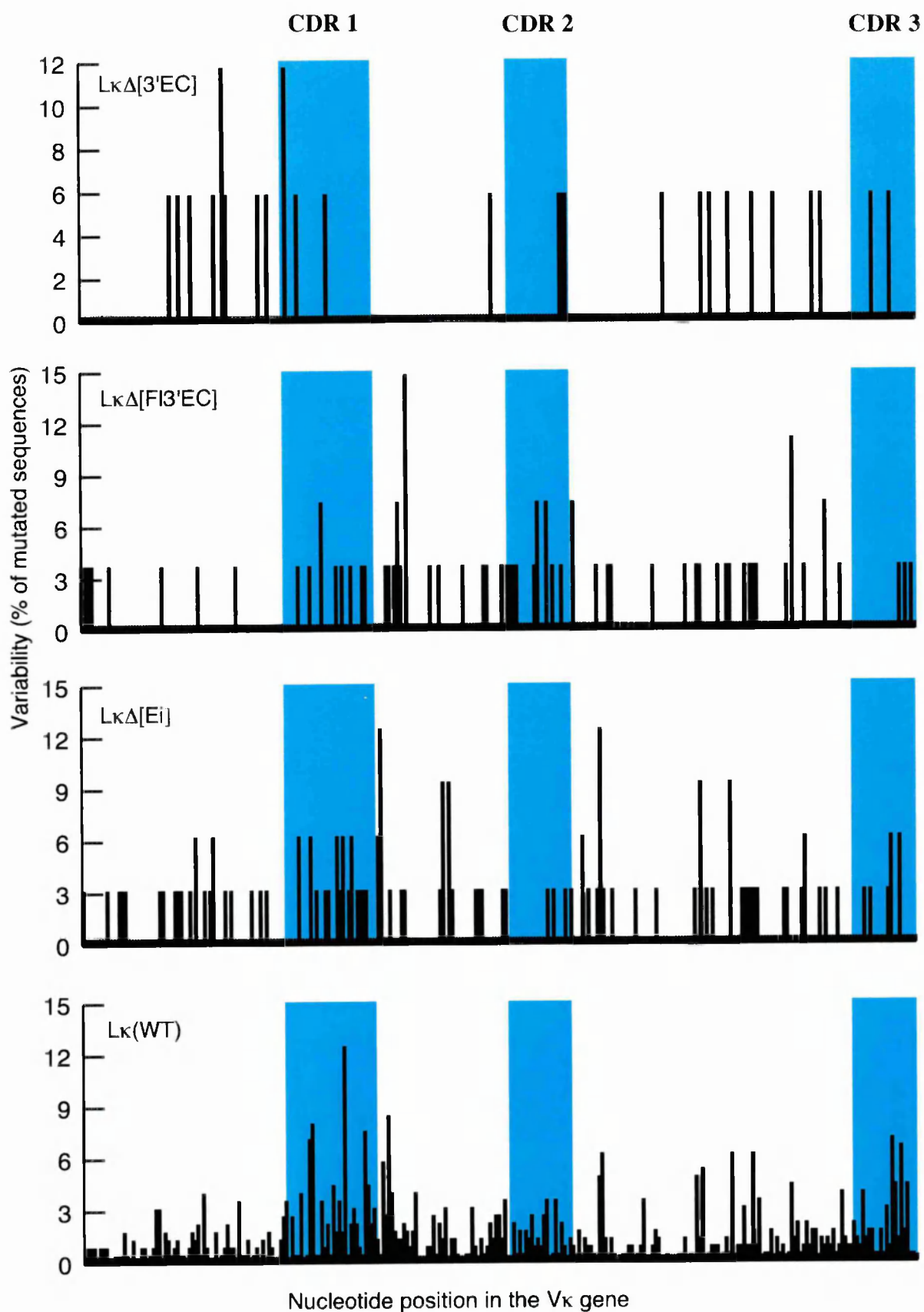
In the case of the LκΔ[3'EC] transgenic line (founder Δ[3'EC] 21), which lacks the 3' enhancer core, a population of germinal centre PNA^{hi} B cells from Peyer's patches was collected from two of the animals (they were identical to the founder by Southern blot analysis): Δ[3'EC] 26 (4 months and 22 days old) and Δ[3'EC] 57 (5 months old) to analyse hypermutation as described in Materials and Methods. The Vκ O_x1 gene segment was amplified using oligonucleotides VKOXBACK and VKOXFOR (Appendix II) and Taq polymerase. JK5 was used as sequencing primer. The PCR error frequency (0.4×10^{-3} per base) was determined by sequencing transgenes from hybridomas of LκNG transgenic line (see Chapter 4).

The sequences with one or more point mutation are shown in the Appendix III D. The sequence analysis of the LκΔ[3'EC] Vκ region revealed that 60% of the sequences were unmutated (see Table 5.4 and Fig. 5.2).

The mutation frequency amongst clones carrying two or more mutation is 12.8×10^{-3} mutations per base pair. This value is only 1.8 times higher than that for the PCR errors (7) and therefore is an intermediate rate between that obtained for Lκ(WT) (18.7) or LκΔ[F13'EC] (14.4) and LκΔ[3'E] (9.5).

The mutation frequency calculated for all clones (2.4) is six times higher than the maximal PCR error, indicating that the removal of the 3' enhancer core does not completely abolish hypermutation. However it causes a big decrease in the hypermutation frequency. Only one clone with 5 mutations was found (as the most strongly mutated). A similar effect on hypermutation is observed when the whole 3' enhancer is removed in LκΔ[E3']. In both transgenic lines the mutation frequency was reduced but not to the same level as in Lκ[MAR] line, suggesting that the machinery is still able to produce mutations at a "basal level".

The distribution of mutations (Fig. 5.2) differs significantly from the one obtained in Lκ (WT) transgenes. Only 11% of the mutated clones carry more than one mutation, which is lower than the proportion obtained in LκΔ[F13'EC] (52%). It is important to point out that one of the two clones carrying four mutations shares the mutations with the clone that carries five, indicating that they might be clonally related. There were also two more pairs which shared one mutation not located in a hotspot. All of them have been included in the calculations. Therefore, there were almost four lineages carrying more than one mutation out of the 101 sequences analysed.



5.3 Discussion

Derivatives of the L κ transgene were used to analyse the effect of the deletion of regulatory sequences on hypermutation. In κ chains, at least two DNA fragments were shown to be required for full hypermutation (Betz et al., 1994). One is located in the J-C intron, including the matrix attachment region and intron enhancer, the other in the 3' enhancer located about 9 kb downstream of the C κ exon. The 132bp core fragment within the 3' enhancer is responsible for up to 52% of the activity of the whole enhancer (Chapter 3). The ability of four different transgenes, lacking either the matrix attachment region, the intron enhancer region, the 3' enhancer core or its flanking regions to undergo somatic hypermutation was analysed in germinal centre cells from the Peyer's patches of the transgenic mice.

The results suggest that the κ intron enhancer is not crucial to achieve high levels of somatic hypermutation since highly mutated clones were found, although at a lower frequency than in L κ . The increased proportion of unmutated clones may reflect the relation between levels of transcription and hypermutation, since (as shown in Chapter 3) the κ transgene expression in the Δ [Ei] mouse sera is lower than in L κ 6 line (which carries the intact construct). Alternatively, this could be due to the presence of transgenic copies which are not able to mutate. This has to be confirmed through analysis of other founders of L $\kappa\Delta$ [Ei].

The sequence analysis of the transgenic lines L $\kappa\Delta$ [F13'EC] (lacking the flanking regions of the 3' enhancer core) and L $\kappa\Delta$ [3'EC] (lacking the 3' enhancer core) revealed that hypermutation was impaired in both of them, but more so when the enhancer core was deleted. The low overall mutation frequency and the fact that few highly mutated clones were found in the transgenic line containing the enhancer core alone suggest that, even though the level of hypermutation is not significantly decreased, the flanking regions around the 3' enhancer core are somehow required for full activity of the mutation machinery.

However, in the L $\kappa\Delta$ [F13'EC] line I found (at least four) transgenes fragments which contain the 5' segment of the transgene (including the V-C κ elements) but lack the 3' segment (see Chapter 3). The incomplete copies are likely to increase the number of clones carrying unmutated V κ sequences. If the results obtained for the L $\kappa\Delta$ [F13'EC] line are affected for this reason, the 3' enhancer core alone may account for the effect of the 3' enhancer on hypermutation. This has to be confirmed through analysis of other founders of L $\kappa\Delta$ [F13'EC].

In addition to the low overall mutation frequency the lack of highly mutated L $\kappa\Delta$ [3'EC] and L $\kappa\Delta$ [E3'] clones is similar. These results point to the 3' enhancer core as the principally responsible for the importance of the 3' enhancer in hypermutation, even though the absence of either does not completely abolish it.

The results further suggest a link between level of transcription and hypermutation. The deletion of the regions flanking the core 3' enhancer reduced the κ transgene expression in the mouse sera 3 to 13 fold relative to the L κ 6 (wild type). Hypermutation was also affected but several highly mutated clones were found. On the other hand the deletion of the 3' core enhancer and the whole enhancer resulted in a larger reduction of the κ transgene expression (see Chapter 3) and hypermutation.

A crucial finding is that the matrix attachment region seems to be more critical, since its deletion abolished hypermutation. These results agree with findings for L $\kappa\Delta$ [Ei/MAR] transgenic line (Betz et al., 1994) harbouring a deletion that includes both the Ei and the MAR. This further suggests that in L $\kappa\Delta$ [Ei/MAR] the DNA element responsible for the fall in hypermutation was the MAR. So far, these results are only based on sequences from a single founder. It is essential to obtain and analyse more founders to verify these observations.

The connection between the level of transcription and hypermutation becomes therefore even more persuasive. The κ enhancers seem to provide levels of transcription necessary for the mutational machinery. The κ 3' enhancer is stronger than the κ intron enhancer (Meyer et al., 1989) and its core is able to drive up to 52% of the activity of the whole enhancer (Chapter 3). Removal of the Ei did not affect hypermutation as much as removal of the 3' enhancer elements. It has been shown that both the κ 3' enhancer and the κ intron enhancer drive transcription in a synergistic way. The sum of the levels of transcription produced by them in independent constructs do not account for the level of transcription observed when they are in the same construct. The same effect is observed here when the hypermutation is analysed in the different constructs carrying deletions of either enhancer. Therefore hypermutation seems to depend on the level of transcription provided by the enhancers.

Only when the MAR was removed the somatic mutation was completely abolished, even though the transgene is still reasonably well transcribed. It seems therefore that the role of MAR in hypermutation is not only linked to transcription.

These findings confirm the idea that transcription is necessary but not sufficient for somatic mutation. It seems that hypermutation depends on transcription but is recruited by different elements. These elements could be localised within the MAR region and/or require the κ locus to be attached to the nuclear matrix for other reasons than transcription, perhaps targeting the chromatin to a specific subnuclear location enriched for the hypermutation machinery.

6.1 The core of the κ 3' enhancer is able to drive expression and hypermutation of the κ light chain but its flanking regions are not essential for either of these processes

The role of the 3' enhancer core in the transcription of the κ chain was studied in stably transfected S107 mouse myeloma cells. These experiments showed that the core of the 3' enhancer was able to drive transgene κ expression by itself. The 3' enhancer core retained up to 30% of the whole enhancer activity. This value increased to 52% when a 5kb fragment was removed between the C κ and the 3' enhancer. This increase however was not observed when the intact enhancer was tested. On the contrary, the expression level obtained was lower when the 3' enhancer was brought closer to the V region. A similar transcriptional suppression effect was reported for the Ig heavy 3' enhancer when it was brought closer to the μ intron enhancer (Mocikat et al., 1995). The flanking regions of the κ 3' enhancer core may be necessary for the enhancers cooperation. Alternatively, this effect may be due to the presence of a silencer in the deleted fragment that acts in conjunction with the 3' enhancer core but not with the full enhancer.

The removal of the 3' enhancer core flanking regions reduces the κ expression in the serum of the Δ [F13'EC] transgenic mice (which carry 2 complete copies of the transgene) compared to the level observed in L κ 6 line (which carries 5 copies). However the expression level is similar to the one observed for the L κ (WT&7M) line, which carries only one intact copy of L κ transgene. Hypermutation was affected but not as much as when the whole 3' enhancer was removed (Betz et al., 1994). The flanking regions of the 3' enhancer core are therefore essential neither for expression nor for hypermutation. The results demonstrate that the κ 3' enhancer core - as the principal element of the 3' enhancer - plays a leading role in regulating the κ gene expression *in vivo* and retains the capacity to drive hypermutation of the immunoglobulin κ gene.

The lack of transcriptional activation by the 3' enhancer core flanking regions was previously described by Meyer et al. (1990) and Pongubala et al. (1991), but it was unknown whether these regions were important for hypermutation. The ability of the flanking regions to drive hypermutation was studied here in transgenic mice.

When the enhancer core was removed the κ expression in the serum of the mice was severely reduced. Discrepancies between the level of transgene κ expression in the serum of the founder and its offspring are probably produced by an upregulation of the κ expression in the offspring, which underwent transgene rearrangements due to the

intrinsic instability of inverted repeat structures (detected as deletions). Hypermutation was severely reduced (but not abolished) in the absence of the enhancer core (in mice with the same DNA configuration as the founder). This indicates that the flanking regions do not retain the ability of targeting hypermutation.

6.2 The deletion of the κ intron enhancer did not abolish expression or hypermutation

The deletion of the intron enhancer produces a large decrease in the transgene expression in stably transfected NS0 cells. The levels of expression were also reduced in the serum of the Δ [Ei] transgenic mice but not annulled as observed after deletion of the 3' enhancer (Betz et al., 1994). Hypermutation was also affected but it did not exclude the generation of highly mutated clones. This result indicates that the intron enhancer is not solely responsible for the severe decrease of hypermutation of the $L\kappa\Delta$ [Ei/MAR] transgene.

6.3 The MAR is essential for hypermutation

The expression levels in NS0 cells stably transfected with Δ [MAR] transgenes and in the serum of the Δ [MAR] transgenic mice were lower than that those observed in $L\kappa 6$ or $L\kappa NG$ lines, but the transgene was still transcribed. However, the Δ [MAR] transgene did not undergo hypermutation. Thus, these results demonstrate that MAR must have been the element responsible for the very low frequency of mutations observed in the $L\kappa\Delta$ [Ei/MAR] transgene.

This finding confirms the idea that transcription is necessary but not sufficient for somatic mutation. It seems that hypermutation depends on transcription but is recruited by other elements. MARs might be important for the anchorage of the gene in a region rich in the factors necessary for the mutational machinery. Alternatively, specific sequences in MAR might be directly implicated in recruiting factors necessary for hypermutation while still requiring active transcription.

The idea that the Ig κ light chain MAR could be involved in other functions besides DNA anchorage, e.g. in B cell specific demethylation and in V κ -J κ rearrangement, is relatively new (Lichtenstein et al., 1994; Xu et al., 1996). The role of MAR in hypermutation may be related to its role in the demethylation of the locus. There are only a few examples of MAR binding proteins (Dickinson et al., 1992; Zong and Scheuermann, 1995; Müller et al., 1996) and therefore it is difficult to speculate which could be the specific factor related to hypermutation.

A nuclear matrix binding protein has been identified (MAR-BP1), which binds to the Ig heavy chain enhancer MARs (Zong and Scheuermann, 1995). Binding of MAR-BP1 to the IgH enhancer MARs requires all four nuclear factor- μ negative regulator (NF- μ NR) binding sites. Thus, binding of MAR-BP1 and NF- μ NR is mutually exclusive. These

results are consistent with a model for cell type-specific regulation (Scheuermann and Chen, 1989) where the NF- μ NR repressor binding to the IgH enhancer MARs prevents nuclear matrix attachment by interfering with MAR-BP1/MAR interaction. In this model an important aspect of the MAR function in B lymphocytes would be to attach the heavy chain locus to the nuclear matrix thereby bringing the gene into regions of the nucleus that contain high concentrations of transcription factors, RNA polymerase, topoisomerases, etc. However, in inappropriate cells (other than B lymphocytes) NF- μ NR would be expressed and would bind to its recognition sites in the MAR flanking the enhancer, thereby preventing nuclear matrix attachment.

This is in agreement with the fact that MARs have been found to stimulate transcription from defined promoters in a number of systems, but only when the test construct is stably integrated (Blazquez et al., 1989; Xu et al, 1989; Klehr et al., 1991; Forrester et al., 1994). This suggests that MARs are important for the anchorage of expressed genes in a region of high transcription factor concentration around the nuclear matrix.

Recently it was shown *in vitro* that a p53 mutant binds the Ig heavy chain MAR element with high affinity. The binding to the MAR is a possible mechanism by which the mutant p53 activates the expression of genes involved in cell proliferation and tumorigenesis (Müller et al., 1996).

6.4 The role of the primary DNA sequence and hairpin loops on hotspots

The role of the primary DNA sequence and hairpin loops in the formation of hotspots was studied in several transgenic mouse lines carrying synthetic Vk regions with altered DNA sequence in the CDR1. This analysis was performed using mice carrying multiple transgenes and this was particularly useful because several transgenes could be analysed in the same mouse and some transgenes served as internal control (in particular for the mutation rate) for the other transgenes.

In order to assess whether the distributions of mutations in the transgenes carrying the mutagenised CDR1 have changed with respect to the distribution in the WT transgenes I collected a database with all the WT clones (239) of Lk and Lk Δ B transgenes sequenced here and previously from Peyer's patch B cells (Chapter 4). The number of mutations at each position along the Vk segment sequenced between the WT database and the mutant was compared. The statistical significance of the differences in the mutability in each position was examined by the Chi-square test to assess the effect of modifications carried by a transgene on hypermutation.

6.4.1 The primary sequence is important for the specific location of hotspots

Not only the second base but also the third is a mutational hotspot when Ser31 (AGT) was changed (to AGC) in 7M. The formation of this new hotspot indicates the importance of the primary DNA sequence for targeting specific nucleotides. This was further supported by the finding that Ser31 as TCA was not a hotspot. The results shown here confirm that the mutational machinery has a preference for AGY encoded Ser and the bias against TCA encoded Ser as targets for hypermutation. The 7M/TCA mutant demonstrates that the codon bias exists and is caused by the nucleotide preferences of the hypermutation process.

6.4.2 The secondary structure has a lesser contribution on hotspots

Hairpin loops do not seem to be important for the definition of which base is going to be targeted. However they may make the 'favoured' nucleotides accessible for hypermutation by generating pauses during transcription that can activate the repair mechanism (see below).

6.5 The downstream position of the 3' enhancer with respect to the target area is not essential for hypermutation

In the Lk(WT&7M) transgenic line both V κ (the WT and the 7M) were mutated at the same level and with the same intrinsic preferences described previously for the same gene. This is in spite of the fact that one of the copies is truncated and does not contain the 3' enhancer. Therefore the 3' enhancer of the intact neighbour copy is able to drive hypermutation from a 5' position with respect to the V-C elements. This result indicates that the downstream position of the 3' enhancer with respect to the V region is not required to achieve high levels of mutations.

6.6 Not all the transgene copies accumulate the same amount and type of mutations

The analysis of somatic mutation in hybridomas derived from transgenic lines made with a mixture of different transgenes made it possible to study in detail how individual transgene copies mutated. This analysis in the LkNG line showed a considerable difference in the number of accumulated mutations in individual transgene copies both in selectable and non selectable transgenes. These results are in keeping with previous observations for selectable transgenes (Rogerson et al., 1991) and demonstrate that it can not be attributed to antigen selection.

The analysis of the aminoacid substitutions in hybridomas from Lk(WT&7M) and LkNG lines that define antibodies with increased affinity suggested that those mutations occur more frequently in some transgenes than in others. The preferential targeting of the replacement mutations associated with the affinity maturation in these copies could be due to a difference in the transcription. Further experiments are required to confirm this

proposal. The mRNA levels of each copy could be studied in some of the L κ (WT&7M) hybridomas and should provide a clear answer to this question.

6.7 The mechanism of somatic mutation

Ig genes have to be actively transcribed in order to be targets for somatic mutation. Unrearranged V κ genes that are not transcribed (Mather and Perry, 1981) are not mutated (Selsing and Storb, 1981; Gorski et al., 1983). However, unrearranged V λ genes that are transcribed (Picard and Schaffner, 1984b) are mutated (Motoyama et al., 1991, 1994; Weiss and Wu, 1987). In addition, the L κ transgene carrying deletion of the 3' enhancer was poorly expressed and hypermutation was severely affected (Betz et al., 1994). The results shown here are in line with these findings and further support the idea that the rate of transcription may be tied to the rate of somatic mutation. It is possible that a specialised repair-like system coupled to transcription could account for the described phenomena.

Amongst the proposed models 'the error-prone DNA synthesis' (Brenner and Milstein, 1966) encompasses the possibility of a direct link between transcription and hypermutation. Several lines of evidence support the idea that the nucleotide excise mechanism is coupled to transcription, e.g. the preferential repair of the transcribed DNA strand (Hanawalt and Mellon, 1993) and the requirements of components of the TFIIF factor, also necessary for transcription initiation, for nucleotide excision repair (Bootsma and Hoeijmakers, 1993). There is indeed evidence that the rate of spontaneous mutagenesis in yeast increases when the transcription rate is increased (Korogodin et al., 1991).

The presence of DNA nicks or specific proteins involved in signalling hypermutation may cause the RNA polymerase to pause, or to enhance the natural tendency of the polymerase to halt. This could induce the recruitment of factors involved in the error-prone DNA repair (Neuberger and Milstein, 1995). This model involves the introduction of mutations (in the transcribed DNA strand) which must not be repaired back to the original sequence. Several of the DNA polymerases involved in DNA repair indeed lack the proofreading 3' -to- 5' exonuclease activity and therefore manifest a higher degree of misincorporation.

This model accounts for most of the aspects discussed here. The κ enhancers seem to provide levels of transcription necessary for the mutational machinery. They also may be responsible for the gene and cell specificity of the hypermutation. The deletion of the E μ or the F13'EC decreases the transcription level and therefore the possibility that hypermutation occurs, but the error-prone DNA repair mechanism is still active because several highly mutated clones were found. The deletion of the 3' enhancer or its core further reduces transgene transcription, and even fewer mutated clones were found.

When the MAR was deleted the hypermutation was abolished. Therefore MAR could be the entry site for the specific proteins required for hypermutation or the factor required for mutations could be loaded onto the initiation complex if the 'MAR binding proteins' are involved. Alternatively MAR may encourage the formation of local or regional single strand nicks (Neuberger and Milstein, 1995).

This model also could explain the fact that the hypermutation domain starts downstream of the promoter. The existence of mutational hotspots is compatible because the rate of delays and errors could depend on the primary DNA sequence and secondary structures. The elongation-termination decision of the transcription machinery is a process that depends on DNA and RNA sequences and on the interaction of the transcription complex with extrinsic protein factors. The hairpin loops could play an important role by inducing transcription stalling through destabilization of the elongation complex (von Hippel and Yager, 1992). The transcription coupled repair could operate in the absence of DNA lesion. A pause site could make the RNA polymerase to trigger a repair reaction. This could lead to high levels of spontaneous mutagenesis in a frequently transcribed gene (Hanawalt, 1994, 1995).

6.8 Conclusions and outlook

In this work I have systematically dissected the *cis*- acting DNA elements involved in the regulation of somatic hypermutation. One of the most important conclusions to be drawn from this study is that the intron enhancer and the 3' enhancer elements may contribute to hypermutation alone through their role in the κ light chain gene expression. In contrast the MAR seems to be necessary for reasons other than transcription.

So far, these results are only based on sequences from a single founder mouse. It is essential to obtain and analyse more founders to verify these observations. One of the conclusions to be drawn from this study is the necessity of alternatives to the use of conventional transgenic mice, e.g. gene targeting.

Somatic hypermutation occurs during a very short time in the life of the B cell. The attempts to purify a large amount of 'hypermutating' cells to study the Ig transcription levels has not succeeded so far. There is therefore an increasing necessity for *in vitro* systems which would allow to answer specific questions which are too complex in the transgenic animals. Recent advances in cell-free assays which reconstitute mammalian DNA nucleotide excision repair should help to reproduce hypermutation *in vitro*. These systems should help to identify specific factors necessary for mutation. They should also simplify the further dissection of the role of the MAR and the enhancer elements in hypermutation and thus improve our understanding of its molecular mechanism. The results shown here point to a direct link between transcription and hypermutation. This implication should be kept in mind when designing those systems because such cell-free

assays should contain all the necessary elements to allow the transgene transcription and the repair mechanism to act simultaneously.

As the role of MAR in hypermutation may be related to its role in the demethylation it would be interesting to study the methylation pattern in all the transgenic lines - principally in Δ [MAR] - to determine the extent at which those processes are related. It would also be interesting to study whether a MAR from another origin could target hypermutation. A LK transgene derivative in which the MAR is replaced by another MAR sequence could answer this question.

The specific downstream location of the 3' enhancer with respect to the promoter and V elements did not seem to be essential for hypermutation. This finding leads one to consider what will be the effect on hypermutation of moving the MAR to a 5' position with respect to the target area.

The work described here also demonstrates the importance of the primary DNA sequence for the formation of hotspots. The use of transgenes carrying V κ region with altered CDR1 has confirmed the intrinsic preferences of the mutational machinery for AGC coded serine over TCA serines. With regard to the secondary structure they may not be essential hotspot defining elements. However they may influence the accessibility of the 'favoured' nucleotide for mutation, e.g. by generating pauses in the transcription process necessary to trigger the error-prone DNA repair. The size of the hairpin loop to be most thermodynamically effective to induce stalling and the position of the consensus sequence with respect to the loop are some of the key elements that need to be defined.

References Cited

- Adachi, Y., Käs, E. and Laemmli, U. K. (1989). Preferential, cooperative binding of DNA topoisomerase II to scaffold-associated regions. *EMBO J.* 8, 3997.
- Adetugbo, K., Milstein, C. and Secher, S. D. (1977). Molecular analysis of spontaneous somatic mutants. *Nature* 265, 299-304.
- Aguilera, R. J., Hope, T. J. and Sakano, H. (1985). Characterization of immunoglobulin enhancer deletions in murine plasmacytomas. *EMBO J.* 4, 3689-3693.
- Allen, D., Cumano, A., Dildrop, R., Kocks, C., Rajewsky, K., Rajewsky, N., Roes, J., Sablitzky, F. and Siekevitz, M. (1987). Timing, genetic requirements and functional consequences of somatic hypermutation during B-cell development. *Immunol. Rev.* 96, 5-22.
- Alt, F. and Baltimore, D. (1982). Joining of immunoglobulin heavy chain gene segments: Implications from a chromosome with evidence of three D-JH fusions. *Proc. Natl. Acad. Sci. USA* 79, 4118-4122.
- Alzari, P. M., Spinelli, S., Mariuzza, R. A., Boulot, G., Poljak, R. J., Jarvis, J. M. and Milstein, C. (1990). Three-dimensional structure determination of an anti-2-phenyloxazolone antibody: the role of somatic mutation and heavy/light chain pairing in the maturation of an immune response. *EMBO J.* 9, 3807-3814.
- Amzel, L. M. and Poljak, R. J. (1979). Three-dimensional structure of immunoglobulins, *Ann. Rev. Biochem.* 48, 961-997.
- Arpin, C., Dechanet, J., Van Kooten, C., Merville, P., Grouard, G., Briere, F., Banchereau, J. and Liu, Y. -J. (1995). Generation of memory B cells and plasma cells in vitro. *Science* 268, 720-722.
- Atchison, M. L. and Perry, R. P. (1987). The role of the κ enhancer and its binding factor NF- κ B in the developmental regulation of κ gene transcription. *Cell* 48, 121-128.

Azuma, T., Motoyama, N., Fields, L. E. and Loh, D. Y. (1993). Mutations of the chloramphenicol acetyl transferase driven by the immunoglobulin promoter and intron enhancer. *Int. Immunol.* 5, 121-130.

Banerji, J., Olson, L. and Schaffner, W. (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* 33, 729-740.

Benezra, R., Davis, R. L., Lockshon, D., Turner, D. L. and Weintraub, H. (1990). The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. *Cell* 61, 49-49.

Berek, C., Griffiths, G. M. and Milstein, C. (1985). Molecular events during maturation of the immune response to oxazolone. *Nature* 316, 412-8.

Berek, C. and Milstein, C. (1987). Mutation drift and repertoire shift in the maturation of the immune response. *Immunol. Rev.* 96, 23-41.

Berek, C., and Milstein, C. (1988). The dynamic nature of the antibody repertoire. *Immunol. Rev.* 105, 5-26.

Bergman, Y., Rice, D., Grosschedl, R. and Baltimore, D. (1984). Two regulatory elements for immunoglobulin kappa light chain gene expression. *Proc. Natl. Acad. Sci. USA* 81, 7041-7045.

Betz, A. G. (1994). Somatic hypermutation of immunoglobulin κ light chain genes. Ph.D. dissertation, Cambridge University, Cambridge, UK.

Betz, A. G., Milstein, C., González-Fernández, A., Pannell, R., Larson, T. and Neuberger, M. S. (1994). Elements regulating somatic hypermutation of an immunoglobulin κ gene: critical role for the intron enhancer/matrix attachment region. *Cell* 77, 239-248.

Betz, A. G., Neuberger, M. S. and Milstein, C. (1993a). Discriminating intrinsic and antigen-selected mutational hotspots in immunoglobulin V genes. *Immunol. today* 14, 405-411.

Betz, A. G., Rada, C., Pannell, R., Milstein, C. and Neuberger, M. S. (1993b). Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hot spots. *Proc. Natl. Acad. Sci. U S A* 90, 2385-2388.

Blasquez, V. C., Hale, M. A., Trevorow, K. W. and Garrard, W. T. (1992). Immunoglobulin κ gene enhancers synergistically activate gene expression but independently determine chromatin structure. *J. Biol. Chem.* 267, 23888-23893.

Blasquez, V. C., Xu, M., Moses, S. C. and Garrard, W. T. (1989). Immunoglobulin κ gene expression after stable integration; I. Role of the intronic MAR and enhancer in plasmacytoma cells. *J. Biol. Chem.* 264, 21183-21189.

Blier, P. R. and Bothwell, A. (1987). A limited number of B cell lineages generates the heterogeneity of a secondary immune response. *J. Immunol.* 139, 3996-4006.

Blier, P. R., and Bothwell, A. (1988). The Immune response to the hapten NP in C57BL/6 mice: Insight into the structure of the B-cell repertoire. *Immunol. Rev.*, 105, 27-43.

Bode, J., Kohwi, Y., Dickinson, L., Joh, T., Klehr, D., Miele, C. and Kohwi-Shigematsu, T. (1992). Biological significance of unwinding capability of nuclear matrix-associating DNAs. *Science* 255, 195-197.

Bootsma, D. and Hoeijmakers, J. H. J. (1993). Engagement with transcription. *Nature* 363, 114-155.

Both, G. W., Taylor, L., Pollard, J. W. and Steele, E. J. (1990). Distribution of mutations around rearranged heavy-chain antibody variable-region genes. *Mol. Cell. Biol.* 10, 5187-5196.

Bothwell, A. L., Paskind, M., Reth, M., Imanishi, K. T., Rajewsky, K. and Baltimore, D. (1981). Heavy chain variable region contribution to the NPb family of antibodies: somatic mutation evident in a gamma 2a variable region. *Cell* 24, 625-637.

Brack, C., and Tonegawa, S. (1977). Variable and constant parts of the immunoglobulin light chain gene of a mouse myeloma cell are 1250 nontranslated bases apart. *Proc. Natl. Acad. Sci. USA* 74, 5652-6656.

Brack, C., Hirama, M., Lenhard-Schuller, R. and Tonegawa, S. (1978). A complete immunoglobulin gene is created by somatic recombination. *Cell* 15, 1-14.

Brenner, S. and Milstein, C. (1966). Origin of antibody variation. *Nature* 211, 242-246.

Brown, M. Stenzel-Poore, M., Stevens, S., Kondoleon, S. K., Ng, J., Baechinger, H. P. and Rittenberg, M. B. (1992). Immunologic memory to phosphocholine keyhole limpet hemocyanin. Recurrent mutations in the λ 1 light chain increases affinity for antigen. *J. Immunol.* 148, 339-346.

Buluwela, L., Forster, A., Boehm, T. and Rabbitts, T. H. (1989). A rapid procedure for colony screening using nylon filters. *Nucleic Acids Res.*, 17(1).

Burnet, F. M. (1957). A modification of Jerne's theory of antibody production using the concept of clonal selection. *Austral. J. Sci.* 20, 67-69.

Casamayor-Palleja, M., Khan, M. and MacLennan, I. C. M. (1995). A subset of CD4⁺ memory T cells contains preformed CD40 ligand that is rapidly but transiently expressed on their surface after activation through the T cell receptor complex. *J. Exp. Med.* 181, 1293-1301.

Chen, J., Young, F., Bottaro, A., Stewart, V., Smith, R. K. and Alt, F. W. (1993). Mutations of the intronic IgH enhancer and its flanking sequences differentially affect accessibility of the JH locus. *EMBO J.* 12, 4635-4645.

Clark, E. A. and Ledbetter, J. A. (1994). How B and T cells talk to each other. *Nature* 367, 425-428.

Clarke, C., Berenson, J., Gorman, J., Boyer, P. D., Crews, S., Siu, G. and Calame, K. (1982). An immunoglobulin promoter region is unaltered by DNA rearrangement and somatic mutation during B-cell development. *Nucleic Acids Res.* 10, 7731-7749.

Cockerill, P.N. (1990). Nuclear matrix attachment occurs in several regions of the IgH locus. *Nucleic Acids Res.* 18, 2643-2648.

Cockerill, P. N. and Garrad, W. T. (1986). Chromosomal loop anchorage of the κ immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. *Cell* 44, 273-282.

Cockerill, P. N., Yuen, M. H. and Garrad, W. T. (1987). The enhancer of the immunoglobulin heavy chain locus is flanked by presumptive chromosomal loop anchorage elements. *J. Biol. Chem.* 262, 5394-5397.

Coffman, R. L. and Cohn, M. (1977). The class of surface immunoglobulin on virgin and memory B lymphocytes. *J. Immunol.* 118, 1806-2815.

Coffman, R. L., Lebman, D. A. and Rothman, P. (1993). Mechanism and regulation of immunoglobulin isotype switching. *Adv. Immunol.* 54, 22270.

Coleclough, C., Perry, R. P., Karjalainen, K. and Weigert, M. (1981). Aberrant rearrangements contribute significantly to the allelic exclusion of immunoglobulin gene expression. *Nature* 290, 372-378.

Collick, A., Drew, J., Penberth, J., Bois, P., Luckett, J., Scaerou, F., Jefferys, A. and Reik, W. (1996). Instability of long inverted repeats within mouse transgenes. *EMBO J.* 15, 1163-1171.

Crews, S., Griffin, J., Huan, H., Calame, K. and Hood, L. (1981). A single VH gene segment encodes the immune response to phosphorylcholine: somatic mutation is correlated with the class of the antibody. *Cell* 25, 59-66.

Cumano, A. and Rajewsky, K. (1985). Structure of primary anti-(4-hydroxy-3-nitrophenyl)acetyl (NP) antibodies in normal and idiotypically suppressed C57BL/6 mice. *Eur. J. Immunol.* 15, 512-520.

Cumano, A. and Rajewsky, K. (1986). Clonal recruitment and somatic mutation in the generation of immunological memory to the hapten NP [published erratum appears in *EMBO J.* (1989) 8, 2444]. *EMBO J.* 5, 2459-2468.

Currie, R. A. and Roeder, R. G. (1989). Identification of an octamer-binding site in the mouse kappa light-chain immunoglobulin enhancer. *Mol. Cell. Biol.* 9, 4239-4247.

Dariavach, P., Williams, G. T., Campbell, K., Pettersson, S. and Neuberger, M. S. (1991). The mouse IgH 3'-enhancer. *Eur. J. Immunol.* 21, 1499-1504.

Delphin, S. and Stavnezer, J. (1995). Regulation of antibody class switching to IgE: characterization of an Il-4-responsive region in the immunoglobulin heavy-chain germline epsilon promoter. *Ann. N. Y. Acad. Sci.* 764, 123-135.

Dickinson, L. A., Joh, T., Kohwi, Y. and Kohwi-Shigematsu, T. (1992). A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell* 70, 631-645.

Dijkwel, P. A. and Hamlin, J. C. (1988) Matrix attachment regions are positioned near replication sites, genes, and an interamplicon junction in the amplified dihydrofolate reductase domain of chinese hamster ovary cells. *Mol. Cell. Biol.* 8. 5398-5409.

Dildrop, R. (1984). A new classification of mouse VH sequences. *Immunol. Today* 5, 85.

Dreyer, W. J. and Bennett, J. C. (1965). The molecular basis of antibody formation: A paradox. *Proc Natl. Acad. Sci. USA* 54, 864-869.

Dreyfus, M., Doyen, N. and Rougeon, F. (1987). The conserved decanucleotide from the immunoglobulin heavy chain promoter induces a very high transcriptional activity in B-cells when introduced into an heterologous promoter. *EMBO J.* 6, 1685-1690.

Early, P., Huang, H., Davis, M., Calame, K. and Hood, L. (1980). An immunoglobulin heavy-chain variable region gene is generated from three segments of DNA: VH, D and JH. *Cell* 19, 981-992.

Edelman, G. M and Gall, W. E. (1969). The antibody problems. *Ann. Rev. Biochem.* 38, 415-466.

Edelman, G. M. and Poulik, M. D. (1961). Studies on structural units of the γ -globulins. *J. Exp. Med.* 113, 861-884.

Ehrlich, P. (1900). On immunity with special reference to cell life. *Proc. Roy. Soc. London* 66, 424-432.

Eisen, H. N. and Siskind, G. W. (1964). Variations in affinities of antibodies during the immune response. *Biochem.* 3, 996-1014.

Ernst, P and Smale, S. T. (1995). Combinatorial regulation of transcription I: general aspects of transcriptional control. *Immunity* 2, 311-319.

Even, J., Griffiths, G. M., Berek, C. and Milstein, C. (1985). Light chain germ-line genes and the immune response to 2-phenyloxazolone. *EMBO J.* 4, 3439-3445.

Ferrier, P., Kripl, B., Blackwell, K., Furley, A. J., Suh, H., Winoto, A., Cook, W., Hood, L., Consantini, F. and Alt, F. (1990). Separate elements control DJ and VDJ rearrangement in a transgenic recombination substrate. *EMBO J.* 9, 117-125.

Feuillard, J., Taylor, D., Casamayor-Palleja, M., Johnson, G. D. and MacLennan, I. C. M. (1995). Isolation and characteristics of tonsil centroblasts with reference to Ig class switching. *Int. Immunol.* 7, 121-130.

Foote, J. and Milstein, C. (1991). Kinetic maturation of an immune response. *Nature* 352, 530-532.

Forrester, W. D., Van Gerderen, C., Jenuwein, T. and Grosschedl, R. (1994). Dependence of enhancer-mediated transcription of the immunoglobulin μ gene on nuclear matrix attachment regions. *Science* 265, 1221-1225.

Freitas, A. A., Rocha, B. and Coutinho, A. (1986). Lymphocyte population kinetics in the mouse. *Immunol. Rev.* 91, 5-37.

Fulton, R. and Van Ness, B (1993). Kappa immunoglobulin promoters and enhancers display developmentally controlled interactions. *Nucleic Acids Res.* 21, 4941-4947.

Fulton, R. and Van Ness, B (1994). Selective synergy of immunoglobulin enhancer elements in B-cell development: a characteristic of kappa light chain enhancers, but not heavy chain enhancers. *Nucleic Acids Res.* 22, 4216-4223.

Galfré, G. and Milstein, C. (1981). Preparation of monoclonal antibodies: strategies and procedures. *Methods in Enzymol.* 73, 3-46.

Gasser, S. M. and Laemmli, U. K. (1986). Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of *D. melanogaster*. *Cell* 46, 521-530.

Gearhart, P. J. and Bogenhagen, D. F. (1983). Clusters of point mutations are found exclusively around rearranged antibody variable genes. *Proc. Natl. Acad. Sci. U S A* 80, 3439-4343.

Gearhart, P. J., Johnson, N. D., Douglas, R. and Hood, L. (1981). IgG antibodies to phosphorylcholine exhibit more diversity than their IgM counterparts. *Nature* 291, 29-34.

Gerster, T., Picard, D. and Schaffner, W. (1986). During B-cell differentiation enhancer activity and transcription rate of immunoglobulin heavy genes are high before mRNA accumulation. *Cell* 45, 45-52.

Gibson, T. (1984) PhD. Thesis, University of Cambridge, UK.

Gillies, S. D., Morrison, S. L., Oi, V. T. and Tonegawa, S. (1983). A tissue specific transcription enhancer element is located in the major intron of the rearranged immunoglobulin heavy chain gene. *Cell* 33, 717-729.

Gimble, J. M. and Max, E. E. (1987). B-cell nuclear proteins binding in vitro to the human immunoglobulin kappa enhancer: localisation by exonuclease protection. *Mol. Cell. Biol.* 7, 15-25.

Goidl, E. A., Paul, W., E., Siskind, G. W. and Benacerraf, B. (1968). The effect of antigen dose and time after immunization on the amount and affinity of anti-hapten antibody. *J. Immunol.* 100, 371-375.

Golding, G. B., Gearhart, P. J. and Glickman, B. W. (1987). Patterns of somatic mutations in immunoglobulin variable genes. *Genetics* 115, 169-76.

González-Fernández, A., Gilmore, D. and Milstein, C. (1994a). Age related decrease in the proportion of germinal center B cells from mouse Peyer's patches is accompanied by an accumulation of somatic mutations in their immunoglobulin genes. *Eur. J. Immunol.*, 24, 2918-2921.

González-Fernández, A., Gupta, S. K., Pannell, R., Neuberger, M. S. and Milstein, C. (1994b). Somatic mutation of immunoglobulin λ chains: A segment of the major intron hypermutates as much as the complementary determining regions. *Proc. Natl. Acad. Sci. USA* 91, 12614-12618.

González-Fernández, A. and Milstein, C. (1993). Analysis of somatic hypermutation in mouse Peyer's Patches using immunoglobulin κ light chain transgenes. *Proc. Natl. Acad. Sci. USA* 90, 9862-9866.

Gorski, J., Rollini, P., and Mach, B. (1983). Somatic mutations of immunoglobulin variable genes are restricted to the rearranged V gene. *Science* 220, 1179-1181.

- Grabar, P., Fauvert, R., Burtin, P. and Hartman, L. (1956). *Rev. Franc. Etude Clin. Biol.* 1, 175.
- Gray, D., MacLennan, I. C. and Lanen P. J. (1986). Virgin B cell recruitment and the life span of memory clones during antibody responses to 2,4-dinitrophenyl hemocyanin. *Eur. J. Immunol* 16, 641-648.
- Griffiths, G. M., Berek, C., Kaartinen, M. and Milstein, C. (1984). Somatic mutation and the maturation of immune response to 2-phenyl oxazolone. *Nature* 312, 271-275.
- Gussow, D. and Clackson, T. (1989). Direct clone characterization from plaques and colonies by the polymerase chain reaction. *Nucleic Acids Res.* 17:4000.
- Hamlyn, P. H., Gait, M. J. and Milstein, C. (1981). Complete sequence of an immunoglobulin mRNA using specific priming and dideoxynucleotide method of RNA sequencing. *Nucleic Acids Res.* 9, 4485-4494.
- Hanawalt, P. C. (1994). Transcription coupled repair and human disease. *Science* 266, 1957-1958.
- Hanawalt, P. C. (1995). DNA repairs comes of age. *Mutation Res.* 336, 101-113.
- Hanawalt, P. C. and Mellon, I. (1993). Stranded in an active gene. *Curr. Biol.* 3, 67-69.
- Hardie, D. L., Johnson, G. D., Khan, M. and MacLennan, I. C. M. (1993). Quantitative analysis of molecules which distinguish functional compartments within germinal centers. *Eur. J. Immunol.* 23, 997-1004.
- Hardy, R. R., Carmack, C. E., Shinton, S. A., Kemp, D. J., and Hayakawa, K. (1991). Resolution and characterization of pro-B and pre-pro-B cells stages in normal mouse bone marrow. *J. Exp. Med.* 173, 1213-1225.
- Heidelberger, M and Pederson, K. O. (1937). The molecular weight of antibodies. *J. Exp. Med.* 65, 393-414.
- Heremans, R. E., (1960). *Les globulines seriques du systeme gamma.* Edditions Arscia S. A., Brussels.

Hieter, P. A., Korsemeyer, S. J., Waldman, T. A. and Leder, P. (1981). Human immunoglobulin kappa light chain genes are deleted or rearranged in lambda-producing B cells. *Nature* 290, 368-372.

Hill, R. L., Delaney, R., Fellow, R. E. Jr. and Lebowitz, H. E. (1966). The evolutionary origins of the immunoglobulins. *Proc. Natl. Acad. Sci. USA* 56, 1762-1769.

Hood, L. E. and Talmadge D. W. (1970). Mechanisms of antibody diversity: Germ line basis for variability. *Science* 168, 325-334.

Hozumi, N. and Tonegawa, S. (1976). Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc. Natl. Acad. Sci. USA* 73, 3628-3632.

Hromas, R., Pauli, U., Marcuzzi, A., Lafrenz, D., Nick, H., Stein, J., Stein, G., and Van Ness, B. (1988). Inducible DNA-protein interaction of the murine kappa immunoglobulin enhancer in intact cells: comparison with in vitro interactions. *Nucleic Acids Res.* 16, 953-967.

Jacob, J., Kelsoe, G., Rajewsky, K. and Weiss, U. (1991). Intracloal generation of antibody mutants in germinal centres. *Nature* 354, 389-392.

Jacob, J. and Kelsoe, G. (1992). In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. II. A common clonal origin for periaarteriolar lymphoid sheath-associated foci and germinal centers. *J. Exp. Med.* 176, 679-687.

Jacob, J., Przylepa, J., Miller, C. and Kelsoe, G. (1993). In situ studies of the primary immune response to (4-hydroxy-3-nitrophenil)acetyl. III. The kinetics of V region mutation and selection in germinal center B cells. *J. Exp. Med.* 178, 1293-1307.

Jenner, E. (1789) An inquiry into the causes and effects of the variolae vaccinae, known by the name of cow-pox. Sampson Low, London.

Jenuwein, T., Forrester, W. C., Qiu, R.-G. and Grosscheld, R. (1993). The immunoglobulin μ enhancer core establishes local factor access in nuclear chromatin independent of transcriptional stimulation. *Genes and Dev.* 7, 2016-2032.

Jerne, N. K. (1951). A study of avidity based on rabbit skin responses to diphtheria toxin antitoxin mixtures. *Acta Pathol. Microbiol. Scand.* 51 (Suppl. 87), 3-183.

Kaartinen, M., Griffiths, G. M., Markham, A. F. and Milstein, C. (1983). mRNA sequences define an unusually restricted IgG response to 2-phenyloxazolone and its early diversification. *Nature* 304, 320-324.

Kaartinen, M., Pelkonen, E., Even, J. and Mäkelä, O. (1988). V genes of the primary antibody response of C57BL/10 mice to the hapten phenyloxazolone. *Eur. J. Immunol.* 18, 1095-10100.

Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L. McCarrey, J., Cedar, H. and Razin, A. (1992). Developmental pattern of gene-specific DNA methylation in the mouse embryo and germline. *Genes Dev.* 6, 705-714.

Kälberg, E., Gray, D. and Leanderson, T. (1993). Analysis of somatic mutation activity in multiple V kappa genes involved in the response to 2-phenyl-5-oxazolone. *Int. Immunol.* 5, 573-581.

Kepler, T., B. and Perelson, A., S. (1993). Cyclic re-entry of germinal center B cells and the efficiency of affinity maturation. *Immunol. Today* 14, 412-415.

Kitamura, D. and Rajewsky, K. (1992). Targeted disruption of mu chain membrane exon causes loss of heavy-chain allelic exclusion. *Nature* 356, 154-156.

Klaus, G. G., Humphrey, J. H., Kunkl, A. and Dongworth, D. W. (1980). The follicular dendritic cell: its role in antigen presentation in the generation of immunological memory. *Immunol. Rev.* 53, 3-28.

Klehr, D., Maass, K. and Bode, J. (1991). Scaffold-attached regions from the human interferon β domain can be used to enhance the stable expression of genes under the control of various promoters. *Biochemistry* 30, 1264-1270.

Klinman, N. R., Rockey, J. H., Frauenberger, G. and Karush, F. (1966). Equine anti-hapten antibody. 3. The comparative properties of gamma G- and gamma A- antibodies. *J. Immunol.* 96, 587-595.

Klein, S., Sblitzky, F. and Radbruch, A. (1994). Deletion of the IgH enhancer does not reduce immunoglobulin heavy chain production of a hybridoma IgD class switch variant. *EMBO J.* 3, 2473-2476.

- Kocks, C. and Rajewsky, K. (1989). Stable expression and somatic hypermutation of antibody V regions in B-cell developmental pathways. *Ann. Rev. Immunol.* 7, 537-559.
- Köhler, G., and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* 256, 495-7.
- Kohwi-Shigematsu, T. and Kohwi, Y. (1990). Torsional stress stabilizes extended base unpairing in suppressor sites flanking immunoglobulin heavy chain enhancer. *Biochemistry* 29, 9551-9560.
- Kolchanov, N. A., Solovyov, V. V. and Rogozin, I. B. (1987). Peculiarities of immunoglobulin gene structures as a basis for somatic mutation emergence. *FEBS Lett.* 214, 87-91.
- Korogodin, V. I., Korogodina, V. L., Fajsz, C., Chepurnoy, A. I., Mikhova-Tsenova, N. and Simonyan, N. V. (1991). On the dependence of spontaneous mutations on the functional state of genes. *Yeast* 7, 105-117.
- Korsmeyer, S. J., Hieter, P. A., Ravetch, J. V., Poplack, D. G., Waldman, T. A. and Leder, P. (1981). Developmental hierarchy of immunoglobulin gene rearrangements in human leukaemic pre-B-cells. *Proc. Natl. Acad. Sci. USA* 78, 7096-7100.
- Kroese, F. G. M., Timens, W. and Nieuwenhuis, P. (1990). Germinal center reaction and B lymphocytes: morphology and function. In *Currents Topics in Pathology: Reaction Pattern of the Lymph Node*. (Berlin: Springer-Verlag), pp. 103-148.
- Kühn, R., Rajewsky, K. and Müller, W. (1991). Generation and analysis of interleukin-4 deficient mice. *Science* 254, 707-710.
- Kunkel, T. A., and Alexander, P. S. (1986). The base substitution fidelity of eukaryotic DNA polymerases. Mismatching frequencies, site preferences, insertion preferences, and base substitution by dislocation. *J. Biol. Chem.* 261, 160-166.
- Küppers, R., Zhao, M., Hasmann, M. L. and Rajewsky, K. (1993). Tracing B cell development in human germinal centres by molecular analysis of single cells picked from histological sections. *EMBO J.* 12, 4955-4967.
- Kurosawa, Y. and Tonegawa, S. (1982). Organization structure and assembly of immunoglobulin heavy chain diversity DNA segments, *J. Exp. Med.* 155, 201-218.

LaBella, F., Sive, H. L., Roeder, R. G. and Heintz, N. (1988). Cell-cycle regulation of a human histone H2b gene is mediated by the H2b subtype-specific consensus element. *Genes Dev.* 2, 32-39.

Landsteiner, K. (1945). The specificity of serological reactions (revised edition). Harvard University Press, Cambridge, Mass.

Leanderson, T., Kallberg, E. and Gray, D. (1992). Expansion, selection and mutation of antigen-specific B cells in germinal centers. *Immunol. Rev.* 126, 47-61.

Lebecque, S. G. and Gearhart, P. J. (1990). Boundaries of somatic mutation in rearranged immunoglobulin genes: 5' boundary is near the promoter, and 3' boundary is approximately 1kb from V(D)J gene. *J. Exp. Med.* 172, 1717-1727.

Lederman, S., Yellin, M. J., Inghirami, G., Lee, J. J., Knowles, D. M., and Chess, L. (1992). Molecular interactions mediating T-B lymphocyte collaboration in human lymphoid follicles: roles of T-B cell activating molecule (5c8 antigen) and CD40 in contact-dependent help. *J. Immunol.* 149, 3817-3826.

Lehrach, H., Diamond, D., Wozney, J. M. and Boedtker, H. (1977). RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. *Biochemistry* 16, 4743-7751.

Lenardo, M., Pierce, J. W. and Baltimore, D. (1987). Protein-binding sites in Ig gene enhancers determine transcriptional activity and inducibility. *Science* 236, 1573-1577.

Lewis, S., Rosenberg, N., Alt, F. and Baltimore, D. (1982). Continuing kappa-gene rearrangement in a cell line transformed by Abelson murine leukemia virus. *Cell* 30, 807-816.

Lichtenstein, M., Keini, G., Cedar, H. and Bergman, Y. (1994). B cell specific demethylation: A novel role for the intronic κ chain enhancer sequence. *Cell* 76, 913-923.

Lieberson, R., Giannini, S. L., Birshtein, B. K. and Eckhardt, L. A. (1991). An enhancer at the 3' end of the mouse immunoglobulin heavy chain locus. *Nucleic Acids Res.* 19, 933-937.

- Liu, Y. J., Johnson, G. D., Gordon, J. and MacLennan, I. C. (1992). Germinal centres in T-cell-dependent antibody response. *Immunol. Today* 13, 17-21.
- Liu, Y. J., Joshua, D. E., Williams, G. T., Smith, C. A., Gordon, J. and MacLennan, I. C. (1989). Mechanism of antigen-driven selection in germinal centres. *Nature* 342, 929-931.
- Liu, Y. J., Malisan, F., de Bouteiller, O., Guret, C., Lebecque, S., Banchereau, J., Mills, F. C., Max, E. E. and Martinez-Valdez, H. (1996). Within germinal centers, isotype switching of immunoglobulin genes occurs after the onset of somatic mutation. *Immunity* 4, 241-250.
- Lozano, F., Rada, C., Jarvis, J. M. and Milstein, C. (1993). Affinity maturation leads to differential expression of multiple copies of a κ light-chain transgene. *Nature* 363, 271-273.
- Luo, Y. and Roeder, R. G. (1995). Cloning, functional characterization, and mechanism of action of the B-cell-specific transcriptional coactivator OCA-B. *Mol. Cell. Biol.* 15, 4115-4124.
- MacLennan, I. C. and Gray, D. (1986). Antigen-driven selection of virgin and memory B cells. *Immunol. Rev.* 91, 61-85.
- MacLennan, I. C., Liu, Y. J. and Johnson, G. D. (1992). Maturation and dispersal of B-cell clones during T cell-dependent antibody responses. *Immunol. Rev.* 126, 143-161.
- Maizels, N. (1989). Might gene conversion be the mechanism of somatic hypermutation of mammalian immunoglobulin genes? *Trends. Genet.* 5, 4-8.
- Mäkelä, O. and Karjalainen, K. (1977). Inherited immunoglobulin idiotypes of the mouse. *Immunol. Rev.* 34, 119-138.
- Malipiero, U. V., Levy, N. S. and Gearhart, P. J. (1987). Somatic mutation in anti-phosphorylcholine antibodies. *Immunol. Rev.* 96, 59-74.
- Manser, T., Wysocki, L. J., Margolies, M. N. and Gefters, M. L. (1987). Evolution of antibody variable region structure during the immune response. *Immunol. Rev.* 96, 141-162.

Manz, J., Denis, K., Witte, O., Brinster, R. and Storb, U. (1988). Feedback inhibition of immunoglobulin gene rearrangement by membrane mu, but not by secreted mu heavy chains. *J. Exp. Med.* 168, 1363-1381.

Mather, E. L. and Perry, R. P. (1981). Transcriptional regulation of immunoglobulin V genes. *Nucleic Acids Res.* 9, 6855-6867.

Mather, E. L. and Perry, R. P. (1983). Methylation status and DNase I sensitivity of immunoglobulin genes: changes associated with rearrangements. *Proc. Natl. Acad. Sci. USA* 80, 4689-4693.

Matthias, P. and Baltimore, D. (1993). The immunoglobulin heavy chain locus contains another B-cell-specific 3' enhancer close to the α constant region. *Mol. Cell. Biol.* 13, 1547-1553.

McCormack, W. T., Tjoelker, L. W. and Thompson, C. B. (1991). Avian B-cell development: generation of an immunoglobulin repertoire by gene conversion. *Ann. Rev. Immunol.* 9, 21-241.

McHeyzer, W. M., Nossal, G. J. and Lalor, P. A. (1991). Molecular characterization of single memory B cells. *Nature* 350, 502-505.

McKean, D., Huppi, K., Bell, M., Staudt, L., Gerhard, W., and Weigert, M. (1984). Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *Proc. Natl. Acad. Sci. U S A* 81, 3180-3184.

Meyer, K. B. and Ireland, J. (1994). Activation of the immunoglobulin κ 3' enhancer in pre-B cells correlates with the suppression of a nuclear factor binding to a sequence flanking the active core. *Nucleic Acids Res.* 22, 1576-1582.

Meyer, K. B. and Neuberger, M. S. (1989). The immunoglobulin κ locus contains a second, stronger B-cell-specific enhancer which is located downstream of the constant region. *EMBO J.* 8, 1959-1964.

Meyer, K. B., Sharpe, M. J., Surani, M. A. and Neuberger, M. S. (1990). The importance of the 3'-enhancer region in immunoglobulin kappa gene expression. *Nucleic Acids Res.* 18, 5609-5615.

Milstein, C., Even, J. and Berek, C. (1986). Molecular events during the onset and maturation of the antibody response. *Biochem. Soc. Symp.* 51, 173-182.

Milstein, C., Even, J., Jarvis, J. M., González-Fernández, A. and Gherardi, E. (1992). Non-random features of the repertoire expressed by the members of one V kappa gene family and of the V-J recombination. *Eur. J. Immunol.* 22, 1627-1634.

Milstein, C. and Rada, C. (1995). The maturation of the antibody response. in "Immunoglobulin genes" Second edition (T. Honjo and F. W. Alt eds.) Academic Press. pp. 57-81.

Mocikat, R., Kardinal, C. and Klobeck, H. G. (1995). Differential interactions between the immunoglobulin heavy chain intron and 3' enhancer. *Eur. J. Immunol.* 25, 3195-3198.

Motoyama, N., Miwa, T., Suzuki, Y., Okada, H. and Azuma, T. (1994). Comparison of somatic mutation frequency among immunoglobulin genes. *J. Exp. Med.* 179, 395-403.

Motoyama, N., Okada, H. and Azuma, T. (1991). Somatic mutation in constant regions of mouse λ 1 light chains. *Proc. Natl. Acad. Sci. USA* 88, 7933-7937.

Mudd, S. (1932). Hypothetical mechanisms of antibody formation. *J. Immunol.* 23, 423-427.

Müller, B. F., Paulsen, D. and Deppert, W. (1996). Specific binding of MAR/SAR DNA-elements by mutant p53. *Oncogene* 12, 1941-1952.

Nelms, K., Hromas, R. and Van Ness, B. (1990). Identification of a second inducible DNA-protein interaction in the kappa immunoglobulin enhancer. *Nucleic Acids Res.* 18, 1037-1043.

Nelson, K. J., Mather, E. L. and Perry, R. P. (1984). Lipopolysaccharide-induced transcription of the kappa immunoglobulin locus occurs on both alleles and is independent of methylation status. *Nucleic Acids Res.* 12, 1911-1923.

Neuberger, M. S. (1983). Expression and regulation of an immunoglobulin heavy chain gene transfected into lymphoid cells. *EMBO J.* 2, 1373-1378.

- Neuberger, M. S. and Milstein, C. (1995). Somatic hypermutation. *Curr. Opin. Immunol.* 7, 248-254.
- Nisonoff, A., Hopper, J. E. and Spring, S. B. (1975). *The antibody molecule*. New York Ac. Press.
- Nossal, G. J. V., Abbot, A., Mitchell, J. and Lamm, Z. (1968). Antigens in immunology. XI. Ultrastructural features of antigen capture in primary and secondary lymphoid follicles. *J. Exp. Med.* 121, 277-290.
- Nussenzweig, M. C., Shaw, A. C., Sinn, E., Danner, D. B., Holmes, K. L., Morse, H. and Leder, P. (1987). Allelic exclusion in transgenic mice that express the membrane form of immunoglobulin mu. *Science* 236, 816-819.
- O'Brien, R. L., Brinster, R. L., and Storb, U. (1987). Somatic hypermutation of an immunoglobulin transgene in kappa transgenic mice. *Nature* 326, 405-409.
- Oettinger, M. A., Schatz, D. G., Gorka, C. and Baltimore, D. (1990). RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 248, 1517-1523.
- Okumura, K., Julius, M. H., Tsu, T., Herzenberg, L. A. and Herzenberg, L. A. (1976). Demonstration that IgG memory is carried by IgG-bearing cells. *Eur. J. Immunol.* 6, 467-472.
- Orlandi, R., Gussow, D. H., Jones, P. T. and Winter, G. (1989). Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. *Proc. Natl. Acad. Sci. U S A* 86, 3833-3837.
- Papavasiliou, F., Jankovich, M., Suh, H. and Nussenzweig, M. C. (1995). The cytoplasmic domains of immunoglobulin (Ig) alpha and Ig beta can independently induce the precursor B cell transition and allelic exclusion. *J. Exp. Med.* 182, 1389-1394.
- Park, K. and Atchison, M. L. (1991). Isolation of a candidate repressor/activator, NF-E1 (YY-1, γ), that binds to the immunoglobulin κ 3' enhancer and the immunoglobulin heavy-chain μ E1 site. *Proc. Natl. Acad. Sci. USA* 88, 9804-9808.

Parslow, T. G., Blair, D. L., Murphy, W. J., and Granner, D. K. (1984). Structure of the 5' ends of immunoglobulin genes: a novel conserved sequence. *Proc. Natl. Acad. Sci. U S A* 81, 2650-2654.

Pascual, V., Liu, Y. -J., Magalski, A., de Bouteiller, O., Banchereau, J. and Capra, J. D. (1994). Analysis of somatic mutation in five B cell subsets of human tonsil. *J. Exp. Med.* 180, 329-339.

Pauling, L. (1940). Theory of structure and process of formation of antibodies. *J. Am. Chem. Soc.* 62, 2643-2657.

Perry, R. P., and Kelley, D. E. (1979). Immunoglobulin messenger RNAs in murine cell lines that have characteristics of immature B lymphocytes. *Cell* 18, 1333-1339.

Peters, A. and Storb, U. (1996). Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity* 4, 57-65.

Pettersson, S., Cook, G. P., Brüggemann, M., Williams, G. T. and Neuberger, M. S. (1990). A second B-cell specific enhancer 3' of the immunoglobulin heavy-chain locus. *Nature* 344, 8165-8168.

Phi-Van and Strätling, W. H. (1988). The matrix attachment regions of the chicken lysozyme gene co-map with the boundaries of the chromatin domain. *EMBO J.* 7, 655-664.

Picard, D. and Schaffner, W. (1984a). A lymphocyte specific enhancer in the mouse immunoglobulin κ gene. *Nature* 307, 80-82.

Picard, D. and Schaffner, W. (1984b). Unrearranged immunoglobulin lambda variable region is transcribed in kappa -producing myelomas. *EMBO J.* 3, 3031-3035.

Pierani, A., Heguy, A., Fujii, H. and Roeder, R. G. (1990). Activation of octamer-containing promoters by either octamer binding transcription factor 1 (OTF-1) or OTF-2 and requirements of an additional B-cell-specific component for optimal transcription of Immunoglobulin promoters. *Mol. Cell. Biol.* 10, 6204-6215.

Pierce, J. W., Gifford, A. M. and Baltimore, D. (1991). Silencing the expression of the immunoglobulin kappa gene in non-B cells. *Mol. Cell. Biol.* 11, 1431-1437.

Pongubala, J. M. R. and Atchison M. L. (1991). Functional characterization of the developmentally controlled immunoglobulin kappa 3' enhancer: regulation by Id, a repressor of Helix-Loop-Helix transcription factors. *Mol. Cell. Biol.* 11, 1040-1047.

Pongubala, J. M. R. and Atchison M. L. (1995). Activating transcription factor 1 and cyclic AMP response element modulator can modulate the activity of the immunoglobulin κ 3' enhancer. *J. Biol. Chem.* 270, 10304-10313.

Pongubala, J. M. R., Nagulapalli, S., Klemsz, M. J., McKercher, S. R., Maki, R. A. and Atchison M. L. (1992). PU.1 recruits a second nuclear factor to a site important for immunoglobulin κ 3' enhancer activity. *Mol. Cell. Biol.* 12, 368-378.

Pongubala, J. M. R., Van Beveren, C., Nagulapalli, S., Klemsz, M. J., McKercher, S. R., Maki, R. A. and Atchison M. L. (1993). Effect of PU.1 phosphorylation on interaction with NF-EM5 and transcriptional activation. *Science* 259, 1622-1625.

Porter, R. R. (1959). The hydrolysis of rabbit γ -globulin and antibodies with crystalline papain. *Biochem. J.* 73, 119-126.

Potter, H., Weir, L. and Leder, P. (1984). Enhancer-dependent expression of human kappa immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. *Proc. Natl. Acad. Sci. USA* 81, 7161-7165.

Probst, H. and Herzog, R. (1985). DNA regions associated with the nuclear matrix of Ehrlich ascites cells expose single-stranded sites after deproteinization. *Eur. J. Biochem.* 146, 167-171.

Queen, C. and Baltimore, D. (1983). Immunoglobulin gene transcription is activated by downstream sequence elements. *Cell* 33, 741-48.

Queen, C. and Stafford, J. (1984). Fine mapping of an immunoglobulin gene activator. *Mol. Cell. Biol.* 4, 1042-1049.

Rada, C., González-Fernández, A., Jarvis, J. M. and Milstein, C. (1994). The 5' boundary of somatic hypermutation in a V κ gene is in the leader intron. *Eur. J. Immunol* 24, 1453-1457.

Rada, C., Gupta, S. K., Gherardi, E. and Milstein, C. (1991). Mutation and selection during the secondary response to 2-phenyloxazolone. *Proc. Natl. Acad. Sci. U S A* 88, 5508-5512.

Reik, W., Williams, G. T., Barton, S., Norris, M., Neuberger, M. S. and Surani, M. A. (1987). Provision of the immunoglobulin heavy chain enhancer downstream of a test gene is sufficient to confer lymphoid-specific expression in transgenic mice. *Eur. J. Immunol.* 17, 465-469.

Reth, M., Petrac, E., Wiese, P., Lobel, L. and Alt, F. W. (1987). Activation of V kappa gene rearrangement in pre-B cells follows the expression of membrane-bound immunoglobulin heavy chains. *EMBO J.* 6, 3299-3305.

Reynaud, C. A., Mackay, C. R., Muller, R. G. and Weill, J. C. (1991). Somatic generation of diversity in a mammalian primary lymphoid organ: the sheep ileal Peyer's patches. *Cell* 64, 995-1005.

Rogerson, B. J. (1994). Mapping the upstream boundary of somatic mutations in rearranged immunoglobulin transgenes and endogenous genes. *Mol. Immunol.* 31, 83-98.

Rogerson, B., Hackett, J. J., Peters, A., Haasch, D. and Storb, U. (1991). Mutation pattern of immunoglobulin transgenes is compatible with a model of somatic hypermutation in which targeting of the mutator is linked to the direction of DNA replication. *EMBO J.* 10, 4331-4341.

Rogozin, I. B. and Kolchanov, N. A. (1992). Somatic hypermutation in immunoglobulin genes. II Influence of neighbouring base sequences on mutagenesis. *Biochem. Biophys. Acta* 1171, 11-18.

Rothenfluh, H. S., Taylor, L., Bothwell, A. L. M, Both, G. W. and Steele, E. J. (1993). Somatic hypermutation in 5' flanking regions of heavy chain antibody variable regions. *Eur. J. Immunol.* 23, 2152-2159.

Rudikoff, S., Pawlita, M. Pumphrey, J. and Heller, M. (1984). Somatic diversification of immunoglobulins. *Proc. Natl. Acad. Sci. USA* 81, 2162-2166.

Sablitzky, F., Weisbaum, D. and Rajewsky, K. (1985). Sequence analysis of non-expressed immunoglobulin heavy chain loci in clonally related, somatically mutated hybridoma cells. *EMBO J.* 4, 3435-3437.

Sakano, H., Kursawa, Y., Weigert, M. and Tonegawa, S. (1981). Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy chain genes. *Nature* 290, 562-565.

Sakano, H., Maki, R., Kursawa, Y., Roeder, W. and Tonegawa, S. (1980). Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy chain genes. *Nature* 286, 676-683.

Saksela, K. and Baltimore, D. (1993). Negative regulation of immunoglobulin kappa light-chain gene transcription by a short sequence homologous to the murine Bi repetitive element. *Mol. Cell. Biol.* 13, 3698-3705.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1990). *Molecular cloning- a laboratory manual*. New York: Cold Spring Harbor Laboratory.

Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. and Roe, B. A. (1980). Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.*, 143(2), 161-78.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U S A* 74, 5463-7.

Schatz, D. G., Oettinger, M. A. and Baltimore, D. (1989). The V(D)J recombination activating gene, RAG-1. *Cell* 59, 1035-1048.

Scheuermann, R. H. and Chen, U. (1989). A developmental-specific factor binds to suppressor sites flanking the immunoglobulin heavy-chain enhancer. *Genes and Dev.* 3, 1255-1266.

Schitteck, B. and Rajewsky, K. (1992). Natural occurrence and origin of somatically mutated memory B cells in mice. *J. Exp. Med.* 176, 427-438.

Schlossman, S. F. and Kabat, E. A. (1962). Specific fraction of a population of antidextran molecules with combining sites of various sizes. *J. Exp. Med.* 116, 535-547.

- Schwager, J., Bürckert, N., Courtet, M. and Du Pasquier, L. (1989). Genetic basis of the antibody repertoire in *Xenopus*: analysis of the VH diversity. *EMBO J.* 8, 2989-3001.
- Seidman, J. G., Max, E. E. and Leder, P. (1979). A κ immunoglobulin gene is formed by site-specific recombination without further somatic mutation. *Nature* 280, 370-375.
- Selsing, E. and Storb, U. (1981). Somatic mutation of immunoglobulin light-chain variable-region genes. *Cell* 25, 47-58.
- Sen, R. and Baltimore, D. (1986). Inducibility of κ immunoglobulin enhancer-binding protein NF- κ B by posttranslational mechanism. *Cell* 47, 921-928.
- Sharpe, M. J., Milstein, C., Jarvis, J. M. and Neuberger, M. S. (1991). Somatic hypermutation of immunoglobulin kappa may depend on sequences 3' of C kappa and occurs on passenger transgenes. *EMBO J.* 10, 2139-2145.
- Sharpe, M. J., Neuberger, M., Pannell, R., Surani, M. A. and Milstein, C. (1990). Lack of somatic mutation in a kappa light chain transgene. *Eur. J. Immunol.* 20, 1379-85.
- Siskind, G. W. and Benacerraf, B. (1969). Cell selection by antigen in the immune response. *Adv. Immunol.* 10, 1-50.
- Southern, P. J. and Berg, P. (1982). Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *J. Mol. Appl. Genet.*, 1(4), 327-41.
- Sperry, A. O., Blasquez, V. C. and Garrard, W. T. (1989). Dysfunction of chromosomal loop attachment sites: illegitimate recombination linked to matrix association regions and topoisomerase II. *Proc. Natl. Acad. Sci. USA* 86, 5497-5501.
- Spratt, B. G., Hedge, P. J., teHeesen, S., Edelman, A. and Broome-Smith, J. K. (1986). Kanamycin-resistant vectors that are analogues of plasmids pUC8, pUC9, pEMBL8 and pEMBL9. *Gene* 41, 337-342.
- Staudt, L. M. and Lenardo, M. J. (1991). Immunoglobulin gene transcription. *Ann. Rev. Immunol.* 9, 373-398.

Steele, E. J. and Pollard, J. W. (1987). Hypothesis: somatic hypermutation by gene conversion via the error prone DNA-RNA-DNA information loop. *Mol. Immunol.* 24, 667-673.

Steiner, L. A. and Eisen, H. N. (1967a). The relative affinity of antibodies synthesized in the secondary response. *J. Exp. Med.* 126, 1185-1205.

Steiner, L. A. and Eisen, H. N. (1967b). Sequential changes in the relative affinity of antibodies synthesized during the immune response. *J. Exp. Med.* 126, 1161-1183.

Stief, A., Winter, D. M., Stratling, W. H. and Sippel, A. E. (1989). A nuclear DNA attachment element mediates elevated and position independent gene activity. *Nature* 341, 343-345.

Storb, U. and Arp, B. (1983). Methylation patterns of immunoglobulin genes in lymphoid cells: correlation of expression and differentiation with under-methylation. *Proc. Natl. Acad. Sci. USA* 80, 6642-6646.

Su, L. K. and Kadesch, T. (1990). The immunoglobulin heavy-chain enhancer functions as the promoter for I mu sterile transcription. *Mol. Cell. Biol.* 10, 2619-2624.

Suresh, M. R., Cuello, A. C. and Milstein, C. (1986) Advantages of bispecific hybridomas in one-step immunocytochemistry and immunoassays. *Proc. Natl. Acad. Sci. USA* 83, 7989-7993.

Tew, J. G. and Mandel, T. E. (1978). The maintenance and regulation of serum antibody levels: evidence indicating a role for antigen retained in lymphoid follicles. *J. Immunol.* 120, 1063-1069.

Tiselius, A. and Kabat, E. A. (1939). An electrophoretic study of immune sera and purified antibody preparations. *J. Exp. Med.* 69, 119-131.

Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature* 302, 575-581.

Umar, A., Schweitzer, P. A., Levy, N. S., Gearhart, J. D. and Gearhart, P. J. (1991). Mutation in a reporter gene depends on proximity to and transcription of immunoglobulin variable genes. *Proc. Natl. Acad. Sci. USA* 88, 4902-4906.

- Valentine R. C. and Green, N. M. (1967). Electron microscopy of an antibody-hapten complex. *J. Mol. Biol.* 27, 615-617.
- von Hippel, P. H. and Yager, T. D. (1992) The elongation-termination decision in transcription. *Science* 255, 809-812.
- von Kries, J. P., Phi- Van, L., Diekmann, S. and Strätling, W. H. (1990). A non-curved chicken lysozyme 5' matrix attachment site is 3' followed by a strongly curved DNA sequence. *Nucleic Acids Res.* 18, 3881-3885.
- Vongsakul, M. (1995). Characterization of immunological memory cells. A minireview. *Asian. Pac. J. Allergy. Immunol.* 13, 75-79.
- Wagner, S. D., Milstein, C. and Neuberger, M. S. (1995). Codon bias targets mutation. *Nature* 376, 732.
- Wabl, M. and Burrows, P. D. (1984). Expression of immunoglobulin heavy chain at high level in the absence of a proposed immunoglobulin enhancer element. *Proc. Natl. Acad. Sci. USA* 81, 2452-2455.
- Weber, J. S., Berry, J., Litwin, S., and Claflin, J. L. (1991). Somatic hypermutation of the JC intron is markedly reduced in unrearranged kappa and H alleles and is unevenly distributed in rearranged alleles. *J. Immunol.* 146, 3218-3226.
- Weigert, M. Cesari, I. M., Yonkovich, S. J. and Cohn, M. (1970). Variability in lambda light chain sequences of mouse antibody. *Nature* 228, 1045-1047.
- Weiss, U. and Rajewsky, K. (1990). The repertoire of somatic antibody mutants accumulating in the memory compartment after primary immunization is restricted through affinity maturation and mirrors that expressed in the secondary response. *J. Exp. Med.* 172, 1681-1689.
- Weiss, U. and Wu, G. E. (1987). Somatic point mutations in unrearranged immunoglobulin gene segments encoding the variable region of lambda light chains. *EMBO J.* 6, 927-932.
- Weiss, U., Zobelein, R. and Rajewsky, K. (1992). Accumulation of somatic mutants in the B cell compartment after primary immunization with a T cell-dependent antigen. *Eur. J. Immunol.* 22, 511-517.

Wirth, T., Staudt, L. and Baltimore, D. (1987). An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. *Nature* 329, 174-178.

World Health Organization (1964). Nomenclature of immunoglobulins and their subunits. *Bull. World Health Organ.* 30, 447.

Wu, T. T. and Kabat, E. A. (1970). An analysis of the sequence of the variable regions of Bence-Jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.* 132, 211-250.

Wysocki, L. J., Geftter, M. L. and Margolies, M. N. (1990). Parallel evolution of antibody variable regions by somatic processes: consecutive shared somatic alterations in VH genes expressed by independently generated hybridomas apparently acquired by point mutation and selection rather than by gene conversion. *J. Exp. Med.* 172, 315-23.

Wysocki, L., Manser, T. and Geftter, M. L. (1986). Somatic evolution of variable region structures during an immune response. *Proc. Natl. Acad. Sci. U S A* 83, 1847-1851.

Xu, M., Hammer, R. E., Blasquez, C. V., Jones, S. L. and Garrard, W. T. (1989). Immunoglobulin κ gene expression after stable integration. II. Role of the intronic MAR and enhancer in transgenic mice. *J. Biol. Chem.* 264, 21190-21195.

Xu, L., Gorham, B., Li, S. C., Bottaro, A., Alt, F. W. and Rothman, P. (1993). Replacement of germ-line ϵ promoter by gene targeting alters control of immunoglobulin heavy chain class switching. *Proc. Natl. Acad. Sci. USA* 90, 3705-3709.

Xu, Y., Davidson, L., Alt, F. W. and Baltimore, D. (1996). Deletion of the Ig κ light chain intronic enhancer/matrix attachment region impairs but not abolish V κ J κ rearrangement. *Immunity* 4, 377-385.

Yancopoulos, G and Alt, F. (1985). Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. *Cell* 40, 271-281.

Yancopoulos, G and Alt, F. (1986). Regulation of the assembly and expression of variable-regions genes. *Ann. Rev. Immunol.* 4, 339-368.

- Yanisch, P. C., Vieira, J. and Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33, 103-19.
- Yélamos, J., Klinx N., Goyenechea, B., Lozano, F., Chui, Y.L., González-Fernández, A., Pannell, R., Neuberger, M.S. and Milstein, C. (1995). Targeting of non-Ig sequences in place of the V segment by somatic hypermutation. *Nature*, 376, 225-29.
- Yisraeli, J. and Szyf, M. (1984). Gene methylation patterns and expression. In "DNA Methylation and Biochemistry and Biological Significance", A. Razin, H. Cedar, A. D. Riggs, eds. (New York: Springer-Verlag) pp. 352-378.
- Zhang, J., Bottaro, A., Li, S., Stewart, V. and Alt, F. W. (1993). A selective defect in IgG2b switching as a result of targeted mutation of the I γ 2b promoter and exon. *EMBO J.* 12, 3529-3537.
- Ziegner, M. Steinhauser, G. and Berek, C. (1994). Development of antibody diversity in single germinal centers: selective expansion of high-affinity variants. *Eur. J. Immunol.* 24, 2393-2400.
- Zong, R.-T. and Scheuermann, R. H. (1995). Mutually exclusive interaction of a novel matrix attachment region binding protein and the NF- μ NR enhancer repressor. *J. Biol. Chem.* 270, 24010-24018.

- Appendix I A** List of oligonucleotides used to generate the CDR1 mutants.
- Appendix I B** List of the selected CDR1 mutants and the mutations they carry.
- Appendix II A** List of other oligonucleotides used.
- Appendix II B** Sequence of the relevant regions in the L κ construct.
- Appendix III A** Sequences with one or more mutations or deletions found in the transgene V κ regions of the germinal centre B cells from Peyer's patches from L $\kappa\Delta$ [Ei] transgenic mice.
- Appendix III B** Sequences with one or more mutations or deletions found in the transgene V κ regions of the germinal centre B cells from Peyer's patches from L $\kappa\Delta$ [MAR] transgenic mouse.
- Appendix III C** Sequences with one or more mutations or deletions found in the transgene V κ regions of the germinal centre B cells from Peyer's patches from L $\kappa\Delta$ [F13'EC] transgenic mice.
- Appendix III D** Sequences with one or more mutations or deletions found in the transgene V κ regions of the germinal centre B cells from Peyer's patches from L $\kappa\Delta$ [3'EC] transgenic mice.
- Appendix III E** Sequences with one or more mutations or deletions found in the L κ transgene V κ regions of the germinal centre B cells from Peyer's patches from L κ (WT&7M) transgenic mice.
- Appendix III F** Sequences with one or more mutations or deletions found in the L κ 7M transgene V κ regions of the germinal centre B cells from Peyer's patches from L κ (WT&7M) transgenic mice.
- Appendix III G** Sequences with one or more mutations or deletions found in the L $\kappa\Delta$ B transgene V κ regions of the germinal centre B cells from Peyer's patches from L κ NG transgenic mice.

Appendix III H Sequences with one or more mutations or deletions found in the LκΔB7M transgene Vκ regions of the germinal centre B cells from Peyer's patches from LκNG transgenic mice.

Appendix III I Sequences with one or more mutations or deletions found in the LκΔB7M/TCA transgene Vκ regions of the germinal centre B cells from Peyer's patches from Lκ7M/TCA&Δ[Li] transgenic mice.

Appendix IV A Sequences of the LκWT and Lκ7M transgenes Vκ regions of anti-phOx hybridomas from Lκ(WT&7M) transgenic mouse.

Appendix IV B Sequences of the LκΔBWT and LκΔB7M transgenes Vκ regions of anti-phOx hybridomas from LκNG transgenic mouse.

Appendix IV C Sequences of the LκΔB7M/TCA transgene Vκ regions of anti-phOx hybridomas from Lκ7M/TCA&Δ[Li] transgenic mouse.

DNA sequence at the CDR 1 of the mutants selected to make transgenic mice.

/----- CDR-1 -----/													
23	24	25	26	27	29	30	31	32	33	34	35	36	
C	S	A	S	S	S	V	S	Y	M	H	W	Y	
5' -C TGC AGT GCC AGC TCA AGT GTA AGT TAC ATG CAC TGG TAC C- 3'													
PstI			SacI		Xho I		NsiI			KpnI			
CDR1M25	C	TGC	AGT	GCG	AGC	TCA	AGC	GTA	AGC	TAT	ATG	CAC	TGG TAC C
CDR1M31	C	TGC	AGT	GCC	AGC	TCG	AGC	GTA	AGC	TAT	ATG	CAC	TGG TAC C
CDR1M28	C	TGC	AGT	GCG	AGC	TCG	AGC	GTA	AGC	TAT	ATG	CAC	TGG TAC C
CDR1M30	C	TGC	AGT	GCC	AGC	TCG	AGC	GTA	AGC	TAT	ATG	CAT	TGG TAC C
CDR1M36	C	TGC	AGT	GCG	AGC	TCA	AGC	GTT	AGC	TAT	ATG	CAC	TGG TAC C
CDR1M7	C	TGC	AGT	GCG	AGC	TCG	AGC	GTT	AGC	TAT	ATG	CAT	TGG TAC C
CDR1M60	C	TGC	AGT	GCC	AGC	TCA	AGT	GTA	TCA	TAC	ATG	CAC	TGG TAC C
CDR1M72	C	TGC	AGT	GCG	AGC	TCG	AGC	GTT	TCA	TAT	ATG	CAT	TGG TAC C

The base changes as well as the new restriction sites introduced appear in red. The restriction sites are underlined.

Oligonucleotides used to generate the CDR 1 mutants derivatives of Lk.

aa/position	C23	S24	A25	S26	S27	S29	V30	S31	Y32	M33	H34	W35	Y36	
CDR1M1	5'	-GT	GCG	AGC	<u>TCG</u>	AGC	GTT	AGC	TAT	ATG	CAT	<u>TGG</u>	<u>TAC</u> - 3'	
CDR1M2	3'	- <u>ACG</u>	<u>TCA</u>	CGC	<u>TCG</u>	<u>AGC</u>	<u>TCG</u>	CAA	TCG	ATA	TAC	GTA	<u>AC</u> - 5'	
CDR1M3	5'	-GT	GCS	AGC	TCR	AGC	GTW	AGC	TAT	ATG	CAY	<u>TGG</u>	<u>TAC</u> - 3'	
CDR1M4	3'	- <u>ACG</u>	<u>TCA</u>	CGS	TCG	AGY	TCG	CAW	TCG	ATA	TAC	GTR	<u>AC</u> - 5'	
CDR1M4					5'	- <u>TCG</u>	<u>AGC</u>	GTT	TCA	TAT	ATG	CAT	<u>TGG</u>	<u>TAC</u> - 3'
CDR1M5					3'	- <u>CG</u>	CAA	AGT	ATA	TAC	GTA	<u>AC</u> - 5'		
CDR1M7	5'	-GT	GCC	AGC	TCA	AGT	GTA	TCA	TAC	ATG	CAC	<u>TGG</u>	<u>TAC</u> - 3'	
CDR1M8	3'	- <u>ACG</u>	<u>TCA</u>	CGG	TCG	AGT	TCA	CAT	AGT	ATG	TAC	GTG	<u>AC</u> - 5'	

aa:amino acid; W: A and T; Y: T and C; R: A and G; S: C and G.

List of oligonucleotides used in this study.

Primers for sequencing

VKOx212: 5' -CTCATTCTCCAGGGTCTGGGTGA- 3'

M13-20 5' -CATTTTGCTGCCGGTCA- 3'

Bgl II

JK5FOR: 5' -CGTTagAtCTCCAGCTTGGTCCCAG- 3'

INJK5: 5' -TAACATGAAAACCTGTGTCTTACACATA- 3'

Oligonucleotides used for PCR amplification

Primers used for direct PCR amplification of the transgene

BamH I

LKFOR: 5' -cgcggAtCCCTTTTCTATCCTGAAGTTCCT- 3'

EcoR I

VKOxBACK: 5' -ccgGggaATTCTCAGCTTCCTGCTAATCA

EcoR I

BEALEADER: 5' -GGGAATTcGAGATCAGAATACAACCAA- 3'

Primer used to screen M13 clones

AFR3: 5' -CCCACTTAAACGTCTAGAGGATCT- 3'

(recognises transgene)

P1400: 5' -ACGTCTAGAAGACCACGCTACCTGCAGTCAGACCC- 3'

(recognises germline)

Primers used to PCR screen the deletion of the intron enhancer

Hind III

MAR1: 5' -TATTAAAAGCTTAATGTATATAATC- 3'

MARBSEQ: 5' -CCTGTCTCTTCCAAGAATACTCTG- 3'

Primers used to PCR screen the deletion of the MAR

IEBSEQ: 5' -CCTTATTTAATTTCTATTGGGAATTAG- 3'

IE5: 5' -TATGGACAGGGCCTTAAGCCAGGGTCTGTATT- 3'

Primers used to amplify the MAR region

MAR1: (described before)

Afl II

MAR2: 5' -TGACTCT_AAAGTAGTTTCAAGAGTT- 3'

Primers used to amplify the DNA segment containing the Ei region

Hind III

IE1: 5' -TTTAT_aAGCTTTTTGTGTTTGACCC- 3'

Not I

RCKN127: 5' -IIIIIIgcGgcCgCGACTGwGGCACCTCCAG- 3'

Primers used to make the 3' enhancer derivatives

Xho I

DINO1: 5' -TTAT_{ctc}GAGTGTCCCATGGACCAA- 3'

Not I

DINO2: 5' -GGAGTGC_{ggCcg}CCAGGCTGTTGGAGG

Primer used for transgene detection in Northern blot

RATCk18: 5' -ATGATGTCCTTATGAACAA- 3'

Primer used for β -actin detection in Northern blot

MOUSEACT: 5' -TTGATGTCACGCACGATTTC- 3'

Appendix II B The sequence of the relevant region in the Lκ construct (shown in the opposite page).

The sequence of the relevant region in the Lκ construct and its derivative is shown. The relevant regions are shown in a different colour for better localisation.

In green is shown the leader region (-238 to -190 and -17 to 1).

The V/J (1 to 324) and the C (starting from 2792) regions are shown in dark blue.

In pink is shown the MAR (1572 to 1936, [Cockerill and Gallard, 1986]).

Within the whole intron enhancer (1864 to 2334, [Queen and Baltimore, 1983; Picard and Schaffner, 1984]), the intron enhancer core (2058 to 2269, [Fulton and Van Ness, 1994]) is indicated in bright blue.

Within the whole 3' enhancer (12680 to 13485, [Meyer and Neuberger, 1989]), the 3' enhancer core (2058 to 2269, [Meyer et al., 1990; Pongubala and Atchison, 1991]) is indicated in cyan.

The promoter elements (octamer: ATTTGCAT, TATA Box: TTATAA and the transcription initiation site: TAATTA [Parslow et al., 1984]) are indicated.

The sequence of all the oligonucleotides used are indicated. The oligo P1400 hybridises with the fragment of the J-C intron of BALB/c deleted in the Lκ construct. The arrow indicates the point of the deletion.

Some relevant restriction endonuclease sites are shown (underlined). The changes incorporated in the oligonucleotides to generate new restriction sites are shown in lower case.

AMAR8324	Q I V L T O S P A I M S A S P G E K V T M T C S A S S V S Y M H W Y C Q X S G T S P X	CAA	ATT	GTT	CTC	ACC	CAG	TCT	CCA	GCA	ATC	ATG	TCT	GCA	TCT	CCA	GCG	GAG	MAC	GTG	ACC	ATG	ACC	TGC	TGC	AGT	GCC	AGC	AGC	TCC	CCC	AAA
AMAR8345	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
AMAR8315	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
AMAR8318	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

		C D R 1																														
		1	5	10	15	20	25	30	35	40	45																					
AMAR8324	R W I Y D T S K L A S G V P A R F S G S G S G T S Y S L T I S S M E A E D A A T Y Y C	AGA	TGG	ATT	TAT	GAC	ACA	TCC	AAA	CTG	GCT	TCT	GCA	GTC	CCT	GAT	GCG	TTC	AGT	GCG	AGT	GCG	TCT	GCG	ACC	TCT	TAC	TAC	TGC	---	---	
AMAR8345	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
AMAR8315	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
AMAR8318	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

		C D R 2																														
		50	55	60	65	70	75	80	85	90	95																					
AMAR8324	R W I Y D T S K L A S G V P A R F S G S G S G T S Y S L T I S S M E A E D A A T Y Y C	AGA	TGG	ATT	TAT	GAC	ACA	TCC	AAA	CTG	GCT	TCT	GCA	GTC	CCT	GAT	GCG	TTC	AGT	GCG	AGT	GCG	TCT	GCG	ACC	TCT	TAC	TAC	TGC	---	---	
AMAR8345	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
AMAR8315	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
AMAR8318	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

		C D R 3																													
		90	95																												
AMAR8324	O O M S S N P	CAG	CAG	TGG	AGT	AGT	MAC	CCA	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
AMAR8345	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
AMAR8315	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
AMAR8318	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Miscellaneous.

Clones 8324, 837, 8332 and 8333 share the same mutation.

		CD R 2										CD R 3									
		50										95									
		55										95									
		60										95									
		65										95									
		70										95									
		75										95									
		80										95									
		85										95									
		90										95									
		95										95									
		100										95									
		105										95									
		110										95									
		115										95									
		120										95									
		125										95									
		130										95									
		135										95									
		140										95									
		145										95									
		150										95									
		155										95									
		160										95									
		165										95									
		170										95									
		175										95									
		180										95									
		185										95									
		190										95									
		195										95									
		200										95									
		205										95									
		210										95									
		215										95									
		220										95									
		225										95									
		230										95									
		235										95									
		240										95									
		245										95									
		250										95									
		255										95									
		260										95									
		265										95									
		270										95									
		275										95									
		280										95									
		285										95									
		290										95									
		295										95									
		300										95									
		305										95									
		310										95									
		315										95									
		320										95									
		325										95									
		330										95									
		335										95									
		340																			

	C D R 3									
	90	91	92	93	94	95	96	97	98	99
DC265	O	O	N	S	S	N	P			
DC266	CAG	CAG	TGG	AGT	AGT	ACC	CCA			
DC267			A--							
DC2614										
DC571										
DC577										
DC579										
DC5710										
DC5711										
DC5713										
DC5714										
DC5716										
DC5720										
DC5721										
DC5724										
DC5728										
DC5729										
DC5730										

Miscellaneous.
Clones 5715 and 267 shared the same mutation.

Sequences of the LxAB transgene Vx region of the germinal center B cells from Peyer's patches from LxNG line (mice 6229 and 6233)

Appendix III G

[illegible]

		C D R 2	
		50	55
R	W I Y D T S R L A	55	60
ACA TGG ATT TAT GAC ACA TCC AAT CTG GCT TCT GGA GTC		60	65
7M/TCAS05177	---	---	---
7M/TCAS05173	---	---	---
7M/TCAS0405	---	---	---
7M/TCAS0401	---	---	---
7M/TCAS04010	---	---	---
7M/TCAS04021	---	---	---
7M/TCAS04022	---	---	---
7M/TCAS04023	---	---	---
7M/TCAS04031	---	---	---
7M/TCAS04032	---	---	---
7M/TCAS04033	---	---	---
7M/TCAS04038	---	---	---
7M/TCAS04040	---	---	---
7M/TCAS04042	---	---	---
7M/TCAS04044	---	---	---
7M/TCAS04045	---	---	---
7M/TCAS04046	---	---	---
7M/TCAS04051	---	---	---
7M/TCAS04057	---	---	---
7M/TCAS0542	---	---	---
7M/TCAS0544	---	---	---
7M/TCAS05413	---	---	---
7M/TCAS05424	---	---	---
7M/TCAS05427	---	---	---
7M/TCAS05432	---	---	---
7M/TCAS05436	---	---	---
7M/TCAS05438	---	---	---
7M/TCAS05446	---	---	---
7M/TCAS05457	---	---	---
7M/TCAS05464	---	---	---
7M/TCAS05469	---	---	---
7M/TCAS05477	---	---	---
7M/TCAS05478	---	---	---
7M/TCAS05484	---	---	---
7M/TCAS05485	---	---	---
7M/TCAS05491	---	---	---
7M/TCAS05497	---	---	---
7M/TCAS054107	---	---	---

Sequences of the LκAB7M/TCA transgene Vκ region of the germinal center B cells from Peyer's patches from Lκ7M/TCABΔ[LJ] line (mice 9051, 9040 and 9054)

	C D R 3									
	90	95								
	O	O	N	S	S	N	P			
	CAG	CAG	TGG	AGT	AGT	AAC	CCA			
7M/TCAB905424	---	---	---	---	---	---	---			
7M/TCAB905427	---	---	---	---	---	---	---			
7M/TCAB905432	---	---	---	---	---	---	---			
7M/TCAB905436	---	---	---	---	---	---	---			
7M/TCAB905438	---	---	---	---	---	---	---			
7M/TCAB905446	---	---	---	---	---	---	---			
7M/TCAB905457	---	---	---	---	---	---	---			
7M/TCAB905464	---	---	---	---	---	---	---			
7M/TCAB905469	---	---	---	---	---	---	---			
7M/TCAB905477	---	---	---	---	---	---	---			
7M/TCAB905484	---	---	---	---	---	---	---			
7M/TCAB905485	---	---	---	---	---	---	---			
7M/TCAB905487	---	---	---	---	---	---	---			
7M/TCAB905497	---	---	---	---	---	---	---			
7M/TCAB9054107	---	---	---	---	---	---	---			

Sequences of the L κ ABWT and L κ AB7M transgenes V κ region from hybridomas of L κ NG line (mice 6233)

Appendix IV B

[illegible]

